(2)

HSD-TR-89-006

EMPIRICAL TESTS OF HYPOTHESES DERIVED FROM A DECISION-THEORETICAL MODEL OF NOISE-INDUCED ANNOYANCE

Sanford Fidell
Laura Silvati
Linda Secrist

October 1989

Final Report for Period September 1987 - July 1989

Noise and Sonic Boom Impact Technology
Human Systems Division
Air Force Systems Command
Brooks Air Force Base, TX 78235-5000

NOTICES

When Government drawings, specifications, or other data are used for any purpose other than in connection with a definitely Government-related procurement, the United States Government incurs no responsibility nor any obligation whatsoever. The fact that the Government may have formulated or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication, or otherwise as in any manner construed, as licensing the holder, or any other person or corporation; or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

The Office of Public Affairs has reviewed this report and it is releasable to the National Technical Information Service (NTIS), where it will be available to the general public, including foreign nationals.

This report has been reviewed and is approved for publication.

ROBERT C. KULL, JR, Capt, USAF
NSBIT Program Manager

FOR THE COMMANDER

MICHAEL G. MACNAUGHTON, COL, USAF
Deputy Commander Development & Acquisition

Please do not request copies of this report from the Human Systems Division. Copies may be obtained from DTIC. Address your request for additional copies to:

Defense Technical Information Center
Cameron Station
Alexandria VA 22301-6145

If your address has changed, if you wish to be removed from our mailing list, or if your organization no longer employs the addressee, please notify HSD/SORT, Brooks AFB TX 78235-5000, to help us maintain a current mailing list.

Copies of this report should not be returned unless return is required by security considerations, contractual obligations, or notice on a specific document.

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION Unclassified	1b. RESTRICTIVE MARKINGS None
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution is unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) BBN Report 6739	5. MONITORING ORGANIZATION REPORT NUMBER(S) HSD-TR-89-006
6a. NAME OF PERFORMING ORGANIZATION BBN Systems & Technologies Corporation 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION HSD/YA-NSBIT
6c. ADDRESS (City, State, and ZIP Code) 21120 Vanowen Street Canoga Park, CA 91303	7b. ADDRESS (City, State, and ZIP Code) Wright-Patterson AFB OHIO 45433-6573
8a. NAME OF FUNDING/SPONSORING ORGANIZATION Noise & Sonic Boom Impact Tech. 8b. OFFICE SYMBOL (If applicable) HSD/YA-NSBIT	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER F33615-86-C-0530
8c. ADDRESS (City, State, and ZIP Code) Wright-Patterson AFB OHIO 45433-6573	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. 63723F PROJECT NO. 3037 TASK NO 02 WORK UNIT ACCESSION NO. 02

11. TITLE (Include Security Classification)
(U) Empirical Tests of Hypotheses Derived From a Decision-Theoretical Model of Noise-Induced Annoyance

12. PERSONAL AUTHOR(S)
Fidell, Sanford, Silvati, Laura and Secrist, Linda

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Final	FROM Sept. 87 TO July 89	88/12/30	62

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	annoyance; environmental impact aircraft noise concentration
05	08		

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

This report describes an empirical study of hypotheses derived from a decision-theoretic model of the annoyance of individual aircraft flyovers. The major issues of interest in this laboratory study were the influences of two internal states of the listener (affective state and concentration on ongoing activity) on the self-reported annoyance of noise intrusions. Support was found for the hypothesis that concentration on ongoing activity influences annoyance judgments independently from the acoustic properties of noise intrusions, and for the hypothesis that the affective state of the observer also influences annoyance judgments independently from the same acoustic properties. However, neither of these influences on annoyance was as large as the main effect of signal level.
The effect of concentration was unambiguous and directly traceable to the experimental manipulation of task difficulty. The experimental manipulations intended to change moods were ineffective, but post hoc analyses of differences between groups of subjects and

(over)

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL Capt. Robert Kull, Program Manager	22b. TELEPHONE (Include Area Code) (513) 255-3376 22c. OFFICE SYMBOL HSD/YA-NSBIT

19. (cont'd) subjects' mood ratings revealed significant differences in annoyance judgments related to alternative definitions of affective state. Keywords; aircraft noise; annoyance.

DTIC COPY INSPECTED 4

Accession For	
NTIS GRA&I	☑
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

Table of Contents

List of Figures

List of Tables

Acknowledgments

The authors are grateful to Dr. David M. Green for discussions of the design and analysis of this experiment; to Ms. Barbara Nicholson for assistance in data collection; to Mr. Suyeo Tomooka for his assistance in acoustic calibration and laboratory preparation; to Mr. Jose DeBarros for assistance in digitizing signals; and to Ms. Elizabeth Fletcher for assistance in production of this report. Thanks are also due to the test subjects for their diligent participation in this study.

Foreword

This report was prepared under Contract F33615-86-C-0530 of the Noise and Sonic Boom Impact Technology (NSBIT) Program. The NSBIT program is conducted by the United States Air Force Systems Command, Human Systems Division under the direction of Captain Robert Kull, Jr.

The work reported herein was undertaken as part of Task Orders 0009, 0013.11, and 0020.4. Mr. Lawrence Finegold was the Technical Monitor.

Executive Summary

An experiment was performed to test hypotheses derived from a decision-theoretic model of the annoyance of individual noise intrusions. The model treats self-reports of annoyance as the product of a rational decision process, in which people discriminate "appropriate" noise exposure (composed of noises which belong to a reference distribution of familiar and tolerated sources) from "inappropriate" noise exposure (composed of noises which belong to a distribution of intrusive sources). The model assigns roles to both acoustic and non-acoustic determinants of annoyance decisions. The hypotheses under test in this study were that the judged annoyance of a noise intrusion is influenced by an observer's affective state and concentration on ongoing activity, independently from the absolute sound pressure level of noise intrusions.

Thirty nine audiometrically screened test subjects were paid to perform a primary (proofreading) task in an anechoic chamber while aircraft flyovers, street traffic noise, and other synthetic noise intrusions were presented for annoyance judgments at sound pressure levels ranging over 30 dB. Each subject served as his own control through exposure to all manipulations of affective state, task difficulty, and noise exposure levels.

In the first round of data collection, 28 subjects were exposed to ten signals presented repeatedly in randomized order at sound pressure levels of 60 to 90dBA while working on a primary task requiring low to moderate levels of concentration. Additionally, these subjects were exposed to conditions intended to create positive and negative affective states through manipulations of certain task parameters. In a second round of data collection, 11 additional subjects were exposed to the same ten signals presented at sound pressure levels ranging from 66 to 96 dBA while working on a primary task requiring low to high levels of concentration. These subjects were also exposed to test conditions intended to influence affective state.

As expected, strong effects of signal presentation level were observed. Annoyance judgments were readily predictable both from A-weighted and audibility-related measures of noise exposure. Support was also found for both of the hypotheses under test. The hypothesis that concentration on ongoing activity influences the apparent annoyance of noise intrusions was supported by effects attributable to the planned manipulations of task difficulty. The hypothesis that affective state influences the apparent annoyance of noise intrusions was supported not by the planned manipulations of mood (which proved largely ineffective), but rather by post hoc analyses of differences among groups of subjects classified by self-rated mood and by earnings of bonus money. Although these findings confirmed the reasonableness of the hypotheses, ethical and other limitations on the ability to manipulate moods of test subjects limited the strength of the observed associations.

Even though additional confirmation of the assumptions of the model could be derived from further laboratory testing (for example, in a between subjects experimental design, or with a more entertaining primary task), it is probably more useful to conduct subsequent tests of the model in the field settings to which its predictions will ultimately be applied. The sort of field tests that could produce evidence of the model's utility would require simultaneous, real time monitoring of noise exposure and annoyance under residential circumstances. Immediate annoyance judgments would be accompanied by longer term measures of affective state and activity at the time of noise intrusions. Miniaturized, computer-based apparatus would have to be developed to record such information.

1. Introduction

This report describes the rationale, conduct, and analysis of a laboratory study of the annoyance of aircraft flyover noise. The experimentation described herein is part of a series of studies undertaken in accordance with the Noise and Sonic Boom Impact Technology's (NSBIT's) planned development of improved means of predicting aircraft noise annoyance, as described by Fidell and Green (1988). The overall purpose of this effort is to develop more accurate and defensible tools for predicting the extensity and intensity of aircraft noise-induced annoyance.

1.1 Background and Rationale for Current Study

For lack of fully satisfactory prediction procedures, aircraft noise-induced annoyance is sometimes dealt with in United States Air Force environmental assessment documents as little more than an inadequately understood, covert mental state. Limitations of methods that are currently available for predicting annoyance have contributed to costly litigation, lengthy delays, and even compromises in training capabilities.

Difficulties in making credible predictions of annoyance are especially pronounced for low altitude, high speed flight along Military Training Routes (MTRs) in sparsely populated areas, and for supersonic flight in Military Operating Areas (MOAs). A large part of the difficulty of predicting annoyance under such conditions is that individual episodes of exposure are generally sporadic; that is, infrequent and aperiodic. Thus, cumulative metrics of annoyance (such as the equivalent level family of exposure - L_{eq}, L_{dn}, L_{dnmr}, etc.) have more face validity in communities near major airports when used as annualized predictors in dosage-response relationships for the prevalence of annoyance than in communities near MOAs and MTRs.

The NSBIT program has adopted a systematic approach to improving methods of predicting aircraft noise annoyance, starting with development of a model of the annoyance of noise exposure. As described by Fidell, Green, Schultz and Pearsons (1988), the approach adopted by NSBIT views annoyance as though it were a rational decision-making process in which both acoustic and nonacoustic factors affect the judged annoyance of noise exposure. The experimentation described in this report is an empirical test of two basic assumptions of this model.

1.2 Overview of Model

A fundamental assumption of the NSBIT model of annoyance is that self-reports of annoyance are determined not only by the acoustic properties of noise intrusions, but also by the state of the listener at the time of the noise intrusion. The following subsections briefly discuss some of the implications of this assumption. Readers interested in additional detail may find it in Fidell et al. (1988).

1.2.1 Acoustic Determinants of Annoyance

Figure 1-1, reproduced from Fidell et al. (1988), asserts that the acoustic basis for judging the annoyance of a noise intrusion is its bandwidth- and duration-adjusted signal-to-noise ratio with respect to some reference distribution. Under residential circumstances of exposure, this reference distribution can be taken to be the interior ambient noise distribution of a dwelling. The annoyance "decision" that people make is whether a noise is more likely to have been associated with the reference distribution of familiar household noises or with a different distribution of intruding noises.

The unit of measurement of bandwidth-adjusted signal-to-noise ratio, a scalar quantity known as d', was developed by Tanner and Birdsall (1958). This quantity has found applications in numerous contexts (cf. Swets, 1964; Green and Swets, 1966), including several related to assessments of the effects of exposure to community noise sources (cf. Fidell and Teffeteller, 1981; Horonjeff, Fidell, Teffeteller, and Green, 1982; and Fidell, Horonjeff, Mills, Baldwin, Teffeteller, and Pearsons, 1985). Fidell, Teffeteller, Horonjeff, and Green (1979), for example, have demonstrated that the annoyance of low absolute level noise intrusions can be predicted from their audibility as measured in units of d'. The acoustic quantity of interest for present purposes is the time integrated audibility of an intruding noise, measured in units of d'-seconds, since annoyance increases with duration of exposure over at least four decades of time (cf. Fidell, Pearsons, Grignetti and Green, 1970).

The physical properties of acoustic signals that influence annoyance decisions are expressed in units of d' for two principal reasons:

- d' measures the ***audibility*** of a noise intrusion (that is, its detectability in a specified ambient noise environment) rather than its absolute sound pressure level; and
- d' measures the audibility of noise intrusions independently from nonacoustic factors (collectively referred to as "response bias") that can affect the self-reported degree of annoyance.

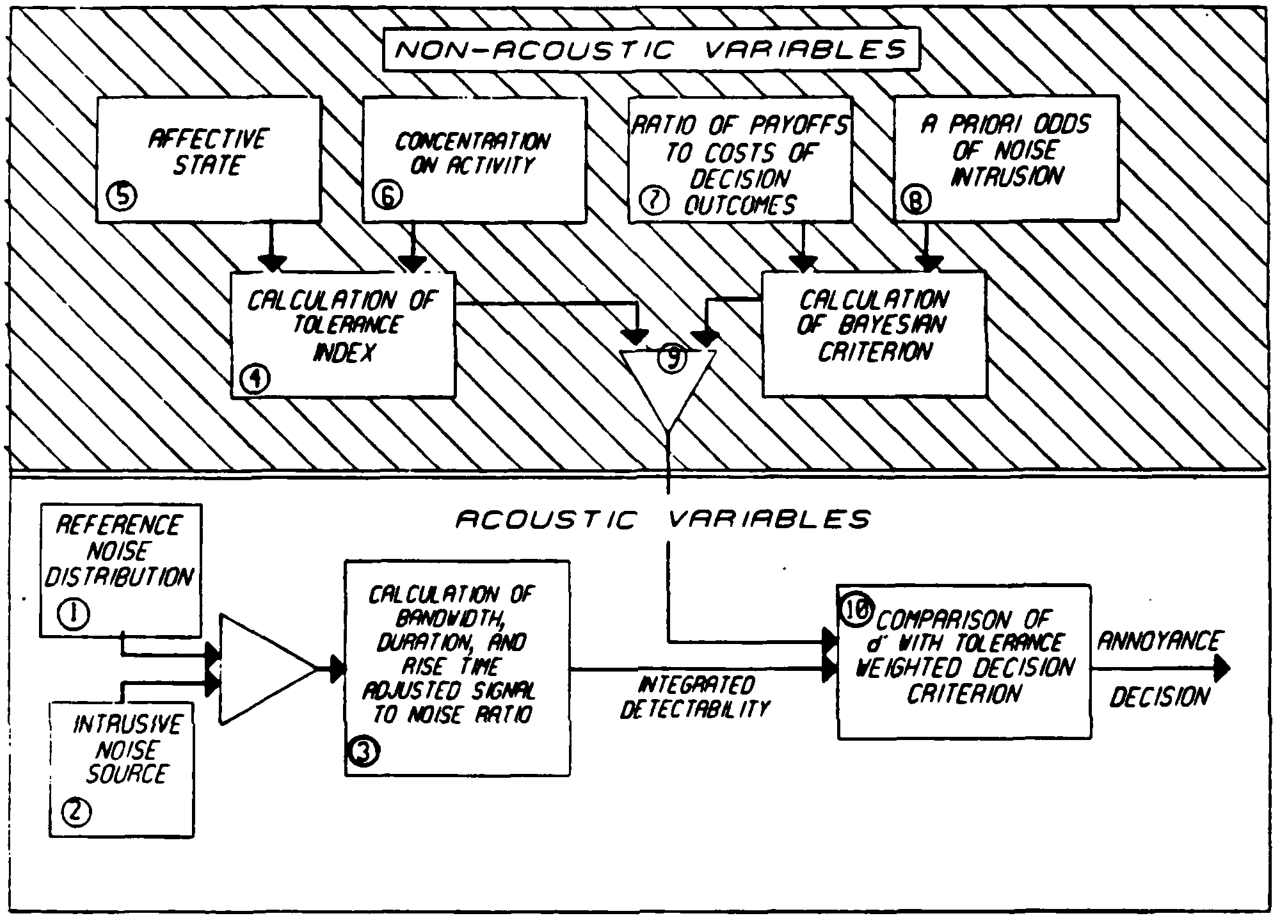


Figure 1-1: Schematic View of Decision-Theoretical Model of Annoyance.

The advantages for present purposes of quantifying noise exposure in terms of its audibility are that:

- Predictions of the acoustically determined annoyance of noise intrusions that are based on measures of signal to noise ratio rather than signal level alone can support predictions of the annoyance of low level noise intrusions (such as those produced by distant aircraft operations in low ambient noise environments), as well as predictions of the annoyance of noise intrusions occurring in non residential environments (occupational, outdoor recreational, etc.) of very different population density; and
- quantification in units of audibility reinforces the distinction between the acoustic and nonacoustic (response bias related) determinants of annoyance.

Note that it is not suggested that scales of audibility supplant currently available acoustic

scales as primary measures of aircraft noise exposure. In fact, for any specific combination of noise source and ambient noise environment, the audibility of a noise intrusion is highly correlated with the absolute A-weighted sound pressure level of the noise intrusion. However, because specification of the relative level of a noise intrusion with respect to the ambient noise environment at the listener's ear permits generalization of predictions of the annoyance of a noise intrusion beyond a particular listening environment, it is useful for environmental impact analysis purposes to be able to estimate the audibility as well as the absolute levels of aircraft noise intrusions.

1.2.2 Nonacoustic Determinants of Annoyance

The two major nonacoustic variables shown in Figure 1-1 as influencing annoyance judgments are the affective state (mood) of the observer and the observer's concentration on ongoing activity. Affective state and concentration on ongoing activity are variables whose quantification has not been commonly attempted in prior studies of the annoyance of noise exposure. There is, therefore, no consensus on standard means for manipulating and measuring these cognitive and emotional factors. Since these factors are not of present interest in themselves, but only as they produce response bias, the most straightforward means of treating them is to define them operationally: that is, in terms of the operations necessary to manipulate them in a specific experimental context. For example, a positive affective state may be operationally defined in the present experimental context as one assumed to exist following some amount of participation in an interesting and rewarding task. A negative affective state may be defined as one prevailing following a comparable amount of participation in a boring or frustrating task.

This well-established practice of operationally defining variables avoids lengthy discussions about psychological interpretations of the independent variables of the current study. The tactic succeeds, however, only to the extent that it can be agreed that the operational definitions adopted are reasonably related to the concepts of interest. An experimental task was, therefore, developed in which conditions could be varied to support plausible arguments about (1) the degree of concentration required for task performance, and (2) the mood of test subjects performing the task.

Effects of concentration on the judged annoyance of noise intrusions were expected to be non-monotonic, as illustrated in Figure 1-2. Just as noise exposure can sometimes facilitate, sometimes impede, and sometimes have no effect on task performance, the judged annoyance of noise intrusions was expected to vary as a function of the level of attention required for task performance. Figure 1-2 suggests that the judged annoyance of noise intrusions is inversely related to concentration on ongoing activity for simple tasks, unrelated to concentration for a

range of tasks of intermediate difficulty, and directly related to concentration on ongoing activity for difficult tasks.

Mood was expected to be directly related to annoyance: the worse a subject's mood, the greater the judged annoyance of a noise intrusion. The major uncertainty was the degree to which experimental manipulations would in fact be effective in changing test subjects' moods.

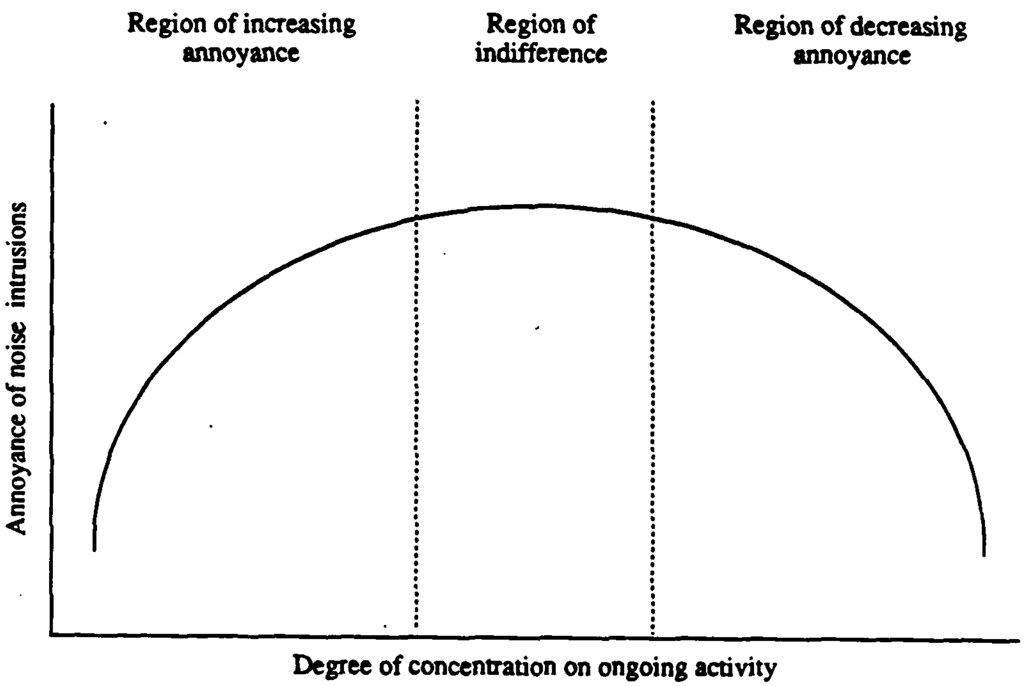


Figure 1-2: Expected Effects of Concentration on Annoyance of Noise Intrusions.

1.3 Logic of Tests of Hypotheses

For illustrative purposes, Fidell et al. (1988) developed spreadsheets to generate quantitative predictions of the decision-theoretic model of annoyance. These spreadsheets embodied arbitrary assumptions about the range of influence of affective state and concentration on ongoing activity on annoyance judgments, based on the premise that these were independent random variables operating over two orders of magnitude (± 10 dB) in equivalent signal level. One purpose of the current experiment was to determine whether these assumptions were reasonable ones.

The hypotheses under test were that test subjects' affective state and concentration on

ongoing activity would influence their judgments of the annoyance of noise intrusions, independently from the absolute sound pressure level of the noise intrusions. A test of these hypotheses required that test subjects make repeated judgments of the annoyance of noise intrusions presented over a range of sound pressure levels while in different affective states and while concentrating to varying degrees on ongoing activity.

It was also possible to collect data useful for testing a number of secondary hypotheses in the course of the current experimentation.[1] One of these hypotheses was that the annoyance of aircraft noise intrusions can be predicted at least as well by measures of audibility as by measures of absolute level. Another hypothesis that could be tested from the same data set was that interference with cognitive task performance can be predicted from measures of the audibility of noise intrusions.

A simple experimental paradigm in which the influences of concentration and mood on the annoyance of noise intrusions could be gauged was one that required test subjects to judge the annoyance of sounds presented to them under acoustically controlled conditions while performing a forced pace primary task. Cost effectiveness and efficiency of data collection and analysis further required that the task be a highly structured one adaptable to automated, repetitive administration.

The primary task also had to be one that could be developed and administered at reasonable expense. This exception implied among other things that the task had to be readily learned by test subjects; that performance had to remain reasonably stable after relatively little practice; and that the task had to be capable of unambiguous scoring. Because it was highly desirable that annoyance judgments be collected from more than one test subject at a time, the task had to be one that could be conducted at a fixed pace, so that multiple subjects' performance could be synchronized. It was also necessary that the task be susceptible to variation in attention demand and satisfaction of performance through simple variations of its parameters.

The task devised to meet these requirements is described in the next Chapter.

[1]Available resources were insufficient to analyze such data.

2. Method

The same experimental paradigm was used with minor variations of task parameters for two rounds of data collection. Descriptions of experimental procedures apply equally to both rounds of data collection except as otherwise noted.

2.1 Physical Conditions of Data Collection

Pairs of test subjects were seated in an anechoic chamber equidistant from a loudspeaker that produced both ambient noise (heard at all times that test subjects were present in the anechoic chamber) and test signals. The purpose for conducting the study under free-field listening conditions was to permit careful control over the spectral content and detectability of the test signals presented for annoyance judgments. The shortest interior dimension (wedge tip to wedge tip) of the anechoic chamber was approximately 4 meters, producing non-reverberant listening conditions at frequencies as low as about 100 Hz.

The conditions of the primary and secondary tasks performed were presented to test subjects via individual computer terminals (viewing screens and keyboards) located at their seats. The terminals were controlled by a MicroVAX II minicomputer, which also controlled all other aspects of the experimentation, including production of test signals from digitized files.

2.2 Experimental Paradigm

A four-way fully factorial within subjects experimental design, as diagrammed in Figure 2-1, was adopted. Test subjects rated the annoyance of 10 signals presented at 6 levels of audibility in the context of 2 levels of concentration and 2 levels of affective state. The 240 (10 x 6 x 2 x 2) resulting conditions were replicated 4 times at the highest signal presentation level, 3 times at the second highest level, and 2 times at the third highest level, as shown in Table 2-1. When combined with 480 trials on which signals were not presented, this yielded a total of 960 trials administered to each test subject.

In Round 1, data were collected from each of 28 test subjects for 2 hr/day, typically over 1 training session and 5 data collection sessions. In Round 2, 960 trials were administered to each of 11 test subjects. Trials were replicated 4 times at the lowest signal presentation level, 3 times at the second lowest level, and 2 times at the third lowest level, as shown in Table 2-2.

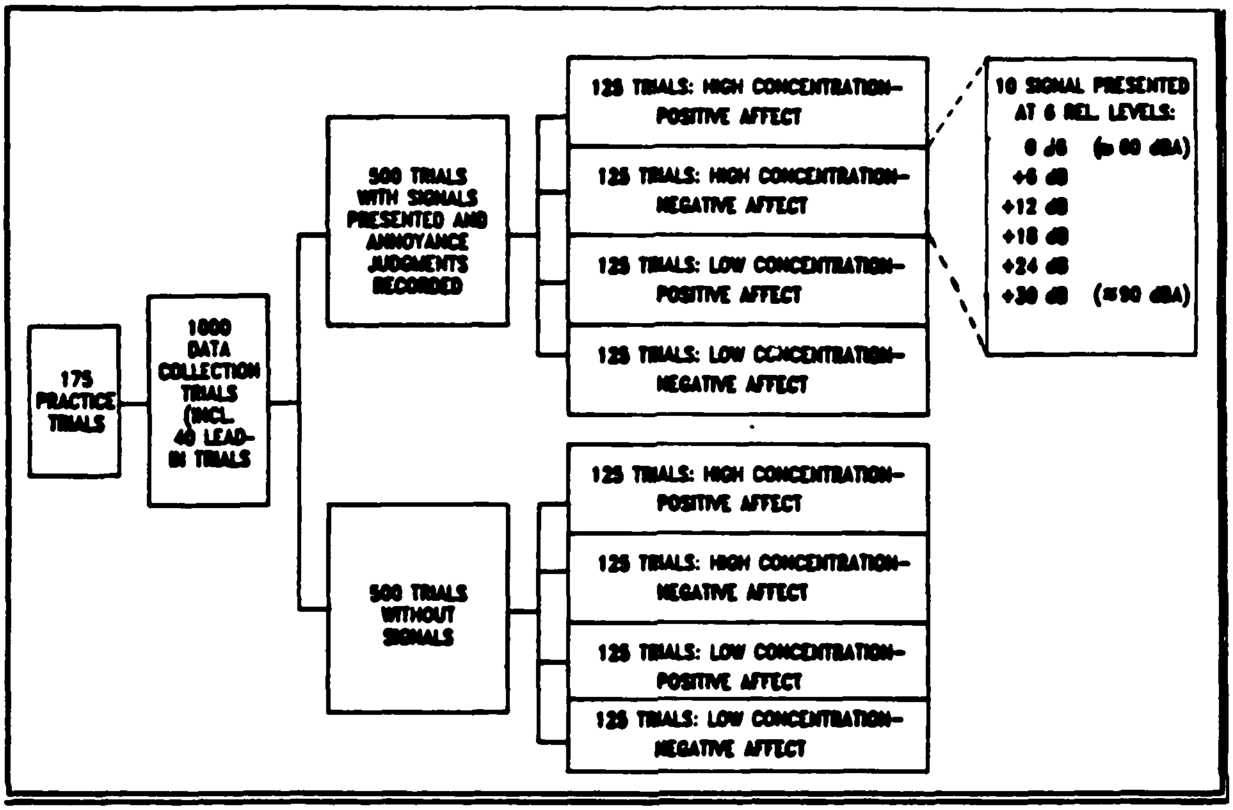


Figure 2-1: Schematic Representation of Experimental Design.

Table 2-1: Numbers of Trials Administered in Each Treatment Condition by Relative Signal Presentation Level for Round 1.

Numbers of Trials					
Relative Signal Level	***High Concentration Positive Affect***	***High Concentration Negative Affect***	***Low Concentration Positive Affect***	***Low Concentration Negative Affect***	***Total # of Trials***
0 dB	10	10	10	10	40
6 dB	10	10	10	10	40
12 dB	10	10	10	10	40
18 dB	20	20	20	20	80
24 dB	30	30	30	30	120
30 dB	40	40	40	40	160
Total	120	120	120	120	480

Table 2-2: Numbers of Trials Administered in Each Treatment Condition by Relative Signal Presentation Level in Round 2.

Numbers of Trials					
Relative Signal Level	***High Concentration Positive Affect***	***High Concentration Negative Affect***	***Low Concentration Positive Affect***	***Low Concentration Negative Affect***	***Total # of Trials***
0 dB	40	40	40	40	160
6 dB	30	30	30	30	120
12 dB	20	20	20	20	80
18 dB	10	10	10	10	40
24 dB	10	10	10	10	40
30 dB	10	10	10	10	40
Total	120	120	120	120	480

Trials were administered in blocks of 25: an initial practice trial to establish the experimental pace, followed by 24 data collection trials. Combinations of task parameters were determined by command files interpreted by the software that administered all experimental conditions, as described in Appendix C. These files described all aspects of trial administration, scoring, and recording of data. Data were stored in files that were automatically analyzed by procedures written in the internal language of RS/1, the statistical software used to conduct most of the analyses reported in the next Chapter.

Two types of trials were administered: Primary Task alone, and Primary plus Secondary Task trials, as described next.

2.3 Primary Task (Character Search) Trials

The primary task required subjects to count the number of occurrences of target characters appearing in lines of non-target characters displayed in a single frame on their terminal screens. The target characters for each trial of the primary task were displayed at the top of the test subjects' screens. The matrix of characters to be searched for target characters next appeared for a variable period of time, and then a demand for a count appeared at the bottom of the screen at the beginning of the response interval. Responses were required within two seconds per target character after the start of the response interval. The response interval was followed

immediately by a feedback interval in which subjects were informed of the accuracy of their answers and any bonuses earned on the trial.

The sequence and duration of the intervals of the primary task and an illustration of the nominal time course signal trials may be seen in Figure 2-2. The starting times of trial intervals subsequent to the display of the primary task's character matrix were contingent upon completion of prior intervals.

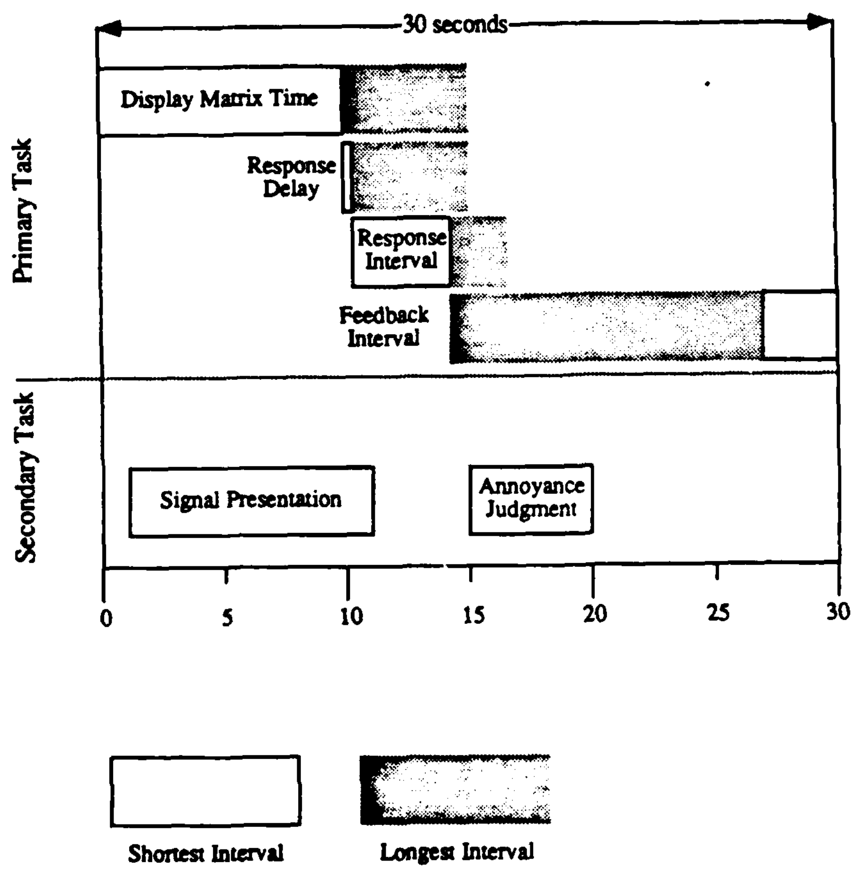


Figure 2-2: Schematic Representation of Order of Presentation of Trial Intervals.

2.3.1 Manipulation of Concentration on Ongoing Activity

Concentration on ongoing activity was manipulated by changing the difficulty of the task. This change was accomplished by varying the number of target characters, number of lines in a frame, number of occurrences of target characters, and the similarity of the target characters to the confuser (background) characters. The degree of concentration demanded by particular combinations of target and confuser characters was quantified in terms of estimated information processing rates. The number of bits per second (b/s) of information processed was proportional to the product of the number of targets and the number of characters divided by the time the frame was displayed.

Task difficulty so measured varied in Round 1 over a range of 5.6 b/s of decision making. An easy task would require only 7.2 to 8.5 b/s of decision making, while a difficult task would require 10.8 to 12.8 b/s of decision making.[2] Task difficulty was held constant over complete blocks of trials. Table 2-3 displays the manipulation of the task parameters that defined the two levels of concentration on ongoing behavior for Round 1.

Task difficulty was altered in Round 2 by increasing the complexity of the search. These more difficult trials required an average of 21.4 b/s of decision making, whereas difficult trials in Round 1 required only 11.8 b/s. The parameter settings for easy trials remained the same for the second set of subjects. Table 2-4 displays the manipulation task parameters that defined the two levels of of concentration of ongoing behavior in Round 2.

2.3.2 Manipulation of Affective State

The experimental manipulation of affective state was intended to alter mood to the greatest extent consistent with ethical considerations and issues of informed consent. Affective state was manipulated by a combination of several of the parameters of the primary task. Cost and payoff contingencies were varied to make the same task either rewarding or punitive. The elapsed time from the end of the search to the start of the response interval was used as a response delay to complicate the reporting of results without affecting the difficulty of the search task. The order in which counts had to be reported for each target character differed from the order in which the target characters were specified. A response lockout was also employed as a manipulation of affective state: the computer was programmed to ignore certain responses, and bonus money was denied as though the responses had not been made in a timely manner. All manipulations of

[2]These figures are nominal ones only; the actual information processing rate demanded by the task depended on the strategy adopted by test subjects for scanning the matrix. It is unlikely that subjects would employ a serial strategy as inefficient as determining whether each character in turn was a member of the target or confuser sets.

Table 2-3: Definition of Conditions for Manipulation of Concentration in Primary Task (Round 1).

Primary Task Conditions					
Concentration Level	*Matrix Size (characters)*	*# of Different Targets*	*Total Occurrences of All Targets*	*Display Duration (seconds)*	*Similarity of Target to Background Characters*
Low*	8 x 8	2	4 to 8	15	low
Low*	6 x 6	2	4 to 12	10	low
High**	8 x 8	3	9 to 14	15	high
High**	6 x 6	3	9 to 16	10	high

*approx. 7.2 - 8.5 b/s
**approx. 10.8 - 12.8 b/s

Table 2-4: Definition of Conditions for Manipulation of Concentration in Primary Task (Round 2).

Primary Task Conditions					
Concentration Level	*Matrix Size (characters)*	*# of Different Targets*	*Total Occurrences of All Targets*	*Display Duration (seconds)*	*Similarity of Target to Background Characters*
Low*	8 x 8	2	4 to 8	15	low
Low*	6 x 6	2	4 to 12	10	low
High**	8 x 8	3	9 to 14	15	high
High**	10 x 10	3	9 to 16	10	high

*approx. 7.2 - 8.5 b/s
**approx. 12.8 - 30 b/s

affective state were held constant over complete blocks of trials. Table 2-5 summarizes the measures taken in an effort to create positive and negative affective states.

Two additional minor manipulations of affect were attempted after completion of data collection from 18 of the first 28 test subjects in Round 1. First, a random schedule of bonus trials (a maximum of 2 trials per block on which triple bonus payments were made for correct responses) was instituted within blocks of positive affect trials. Second, in an effort to minimize boredom during positive affect blocks, humorous one-line messages were displayed during the the feedback intervals in which subjects had no tasks to perform.

Manipulations of affective state in Round 2 were similar to those of Round 1, except for an increase in payoff contingencies. Subjects in Round 2 earned twice as much money on positive affective trials as those in Round 1.

Table 2-5: Description of Conditions for Manipluation of Affective State.

Description of Conditions				
State	*Response Delay Interval*	*Number of Response Lockout Trials per Block*	*Order of Reporting*	*Payoff*
positive	short	0	consistent	high
negative	random	2	random	low

2.4 Secondary Task (Annoyance Judgment) Trials

The secondary task required subjects to judge the annoyance of acoustic signals presented on an unpredictable basis. Procedures for producing and recording these annoyance judgments were arranged to interfere minimally with conduct of the primary task, so that the elicitation of annoyance judgments was not itself unduly annoying. Presentations of sounds for annoyance judgments were synchronous with the matrix display interval of the primary task, with signal presentations less than or equal in length to the display interval.

The annoyance judgments elicited in each of five annoyance category (Extremely, Very, Moderately, Slightly, Not at All) were collected and stored in a table along with the trial parameters defining the subjects' concentration level, affective state, signal presentation level, and signal number.

2.4.1 Test Signals

Several criteria governed selection of test signals.

- First, since a major goal of the NSBIT program is to develop tools which will improve the ability of environmental planners to predict the annoyance of high speed, low altitude aircraft flyover noise, it was important that appropriate aircraft noise signatures be well represented among the test signals.
- Second, because the present experimentation was intended as a test of hypotheses derived from a decision-theoretic view of annoyance, it was also important that the test signals provide an adequate test of the ability to predict annoyance from signal detectability.
- Third, it was important to select signals which could demonstrate the generality of the theoretical predictions for signals other than aircraft flyover noise.
- Fourth, anticipated efforts to link the findings of the current study with those of other NSBIT-sponsored research required collection of annoyance judgments for aircraft noise signature resembling those common in airport environs in the 1960s and 1970s.

Since the spectra of flyover noise signatures were fairly similar to each other, two synthetic signals were created with spectra that differed maximally from one another in their detectability in the ambient noise environment created in the anechoic chamber. These signals were composed of shaped bands of noise whose peak detectabilities occurred in different spectral regions in the mid- and high-frequency spectral regions. Appendix D plots and tabulates the spectra of these test signals.

Finally, since it was also important to demonstrate that hypotheses derived from the decision-theoretic approach to modeling annoyance were not specific to aircraft noise, but were more generally applicable to predicting the annoyance of community noise sources in general, 2 non-aircraft transportation noises were also included. Both noises were vehicle passbys: 1 truck and 1 motorcycle.

The ambient noise environment for the conduct of all testing was selected to approximate a typical interior residential noise environment. Shaped Gaussian noise with an overall spectrum close to PNC-40 was therefore present at all times that test subjects were in the anechoic chamber. The absolute level of the PNC-40 spectrum, about 48 dB(A), sufficed to permit presentation of signals as much as 40 dB higher in A-level that still did not exceed 90 dB(A).

2.4.2 Signal Presentation Levels

Signals were presented for annoyance judgments at six presentation levels 6 dB apart at peak sound pressure levels ranging from 60 dB(A) to 90 dB(A) in Round 1, and from 66 dB(A) to 96 dB(A) in Round 2. The range of audibility corresponding to the six signal presentation levels was 42 dB.[3] Audibility was quantified as bandwidth-adjusted signal to noise ratio in the 24 one-third octave bands from 50 Hz to 10 kHz as:

$$d' = \eta \ S/N \ W^{.5}$$

where η is a parameter that reflects the efficiency of the detector with respect to an ideal energy detector, S is the signal level in a one third octave band, N is the ambient noise level in the same one-third octave band, and W is the width of the detector's input filter. The value of η is generally taken to be 0.4 for human observers. The bandwidths of the hypothetical first stage human auditory filter are narrower than a one-third octave band at frequencies above about 250 Hz but considerably wider at lower frequencies (Fidell, Horonjeff, Teffeteller, and Green, 1982).

2.5 Data Collection Conditions

Data were collected simultaneously but independently from pairs of subjects seated equidistantly from the loudspeaker in the anechoic chamber. One subject was always in a negative affective state while the other was in a positive affective state. To maintain time synchrony between trial intervals for the two subjects working on different primary tasks, the durations of certain trial intervals were of compensating durations so that the completion of trial occurred simultaneously for the two subjects. In the absence of an annoyance judgment (that is, during a trial without an acoustic signal presented for judgment) additional time was added to the feedback interval. Table 2-6 displays the lengths of trial intervals needed to maintain synchrony.

[3]Signals of equal A-weighted sound pressure level are not necessarily equally audible in either the same or different ambient noise environments.

Table 2-6: Trial Interval Durations in all Treatment Conditions.

Affective State and Concentration Level				
Trial Intervals	***Positive Affect*** *Low Concentration (seconds)*	***Positive Affect*** *High Concentration (seconds)*	***Negative Affect*** *Low Concentration (seconds)*	***Negative Affect*** *High Concentration (seconds)*
Display Duration	10 or 15	10 or 15	10 or 15	10 or 15
Response Delay	0	0	2 or 5	1 or 5
Response Interval*	4	6	4	6
Annoyance Judgment**	0 or 5	0 or 5	0 or 5	0 or 5
Feedback Interval	6 - 16	4 - 14	4 - 11	3 - 9
Total Trial Time	30	30	30	30

*Dependent on number of targets

**Dependent on occurrence of noise intrusion

3. Results

This section describes the results of analyses of the major dependent variables of this study: annoyance judgments and self-reports of mood for two independent data collection sessions. Descriptions of findings apply equally to both rounds of data collection except as otherwise noted.

3.1 Descriptive Analyses of Annoyance Judgments

A total of 26,880 trials was administered in Round 1 and 10,560 trials in Round 2. Annoyance judgments of the noise intrusions that occurred on half (13,440 and 5,280, respectively) of these trials are analyzed in this section.

3.1.1 Relationship Between Signal Presentation Level and Annoyance Judgments

Responses made in each of the 5 categories of the annoyance scale (Not at All, Slightly, Moderately, Very, and Extremely) were recorded as arbitrary numeric codes. Analyses of annoyance judgments so represented are restricted to nominal (or at best ordinal) scale procedures such as frequency tabulations and histograms. Even at this level of analysis, a strong relationship between signal presentation level and judged annoyance is apparent. Figure 3-1 plots the numbers of highly annoyed ("Extremely Annoyed" and "Very Annoyed") judgments as a function of signal level in both rounds of data collection. The error bars represent 95% confidence intervals for each plotted point. Figure 3-2 combines the annoyance judgments on a relative basis for the two rounds of data collection.

3.1.2 Reliability of Annoyance Judgments

The availability of repeated measurements of the annoyance of signals of the same presentation level under the same conditions of concentration and affective state permitted examination of the reliability with which subjects were able to make annoyance judgments. Annoyance judgments were expressed for this purpose in terms of standardized scores derived from individual mean annoyance ratings for each test signal. These scores were calculated for each subject as signed differences between each of the repeated annoyance ratings and the mean annoyance rating, divided by the standard deviation of the subject's distribution of annoyance ratings.

No significant differences were observed within subjects among these standard scores. Further, none of the scores was significantly different from zero, it was concluded that the set of annoyance judgments as a whole was a highly reliable one.

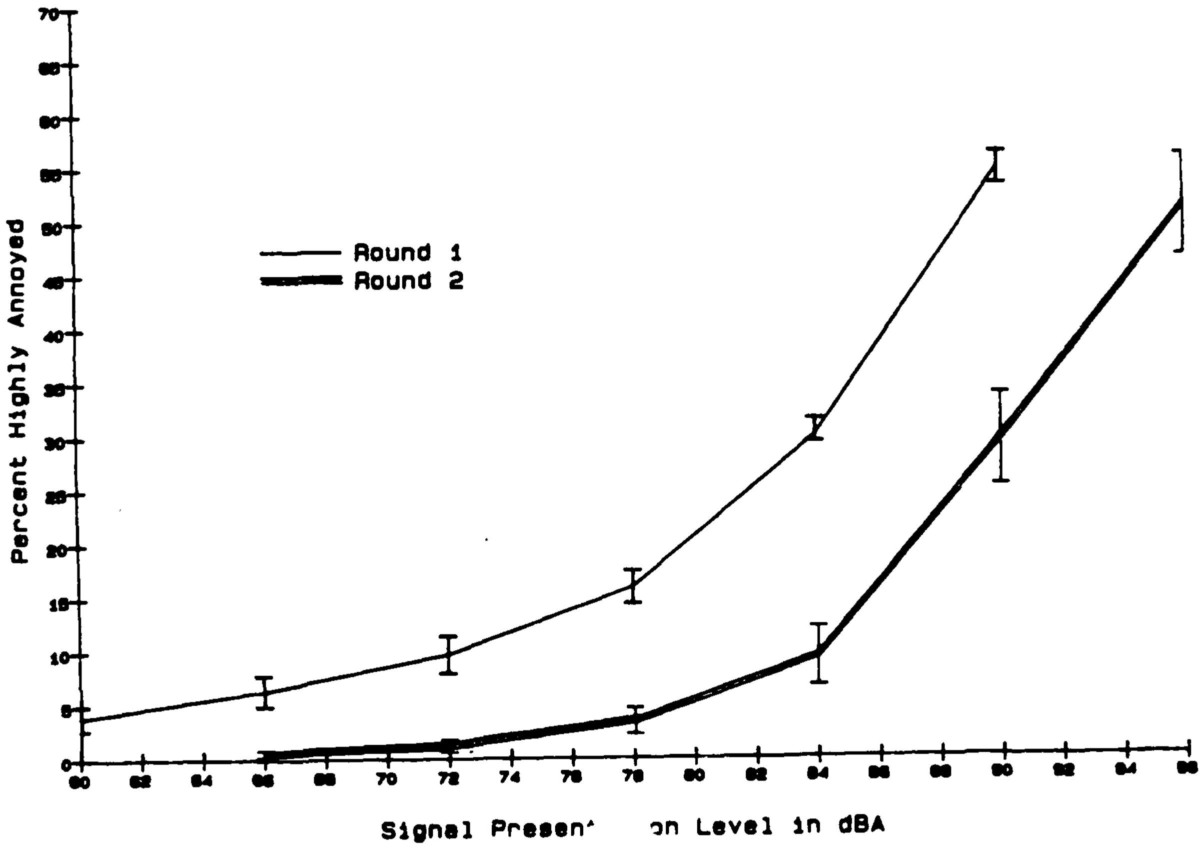


Figure 3-1: Relationship Between Signal Presentation Level and Percentage of Highly Annoyed Responses for all Signals, Conditions and Subjects.

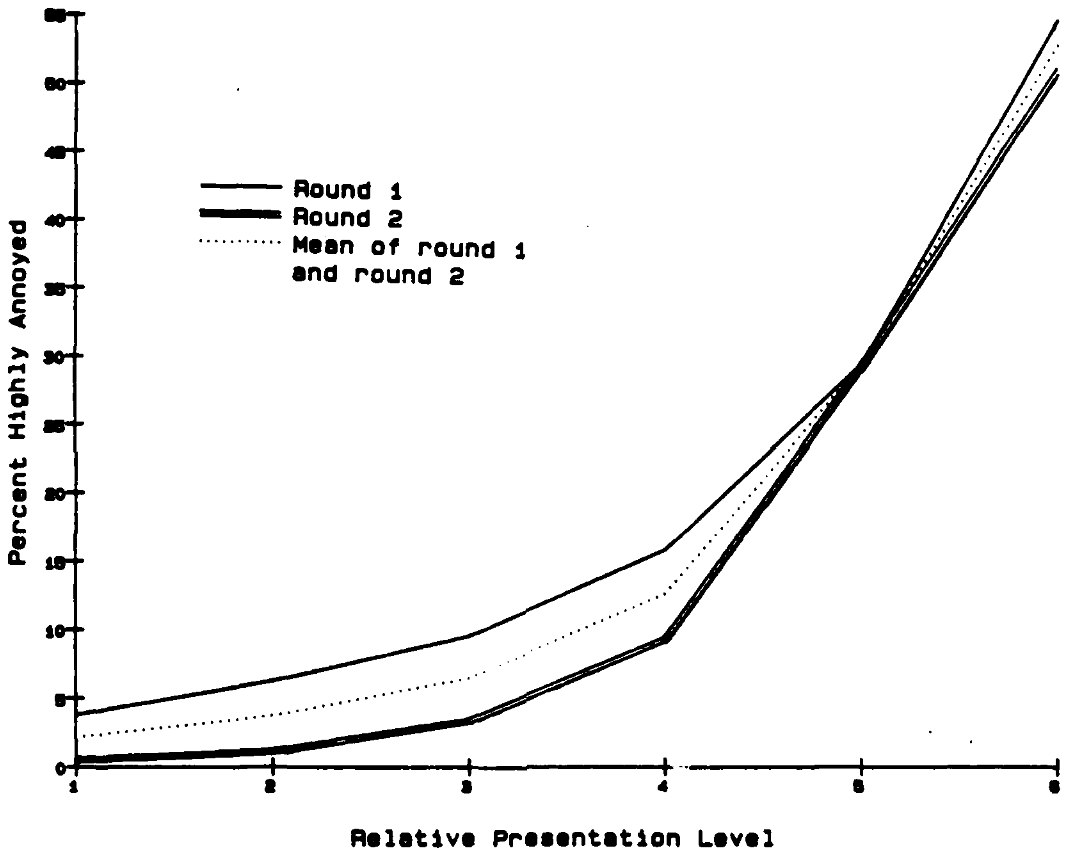


Figure 3-2: Combined Relationship Between Signal Presentation Level and Percentage of Highly Annoyed Responses for all Signals, Conditions and Subjects (Rounds 1 and 2).

3.1.3 Effects of Concentration on Annoyance

An analysis of the effects of concentration on annoyance (collapsed over all treatment conditions) revealed a small difference in the number of highly annoyed judgments in the two concentration conditions in both rounds of data collection, as seen in Figures 3-3 and 3-4.

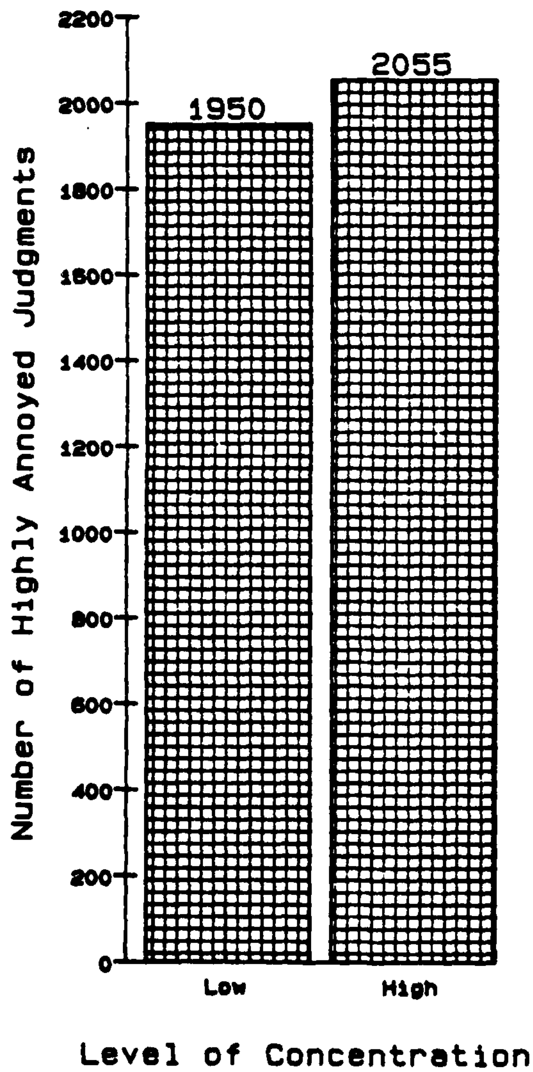


Figure 3-3: Number of Highly Annoyed Judgments for Low and High Concentration Conditions for all Signals, and Levels (Round 1).

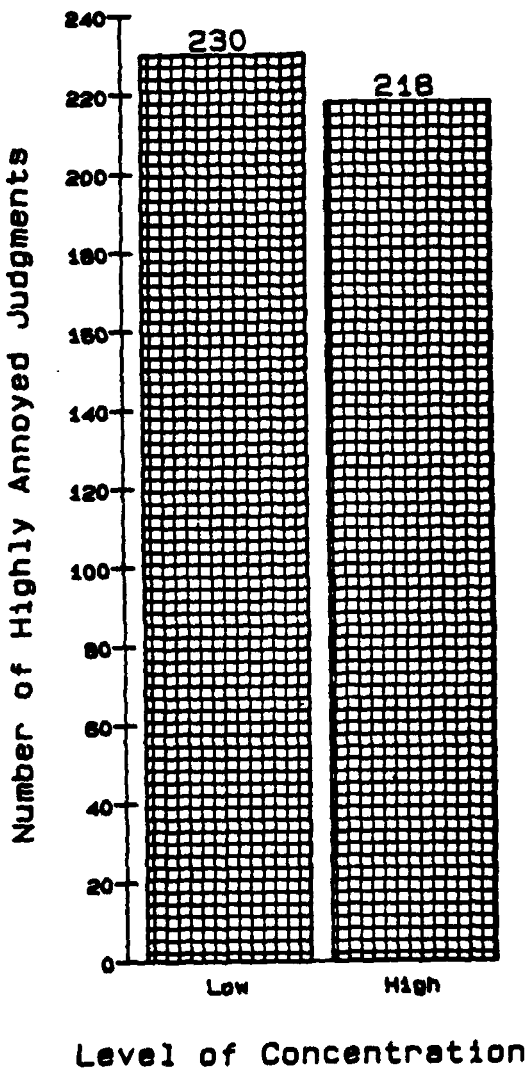


Figure 3-4: Number of Highly Annoyed Judgments for Low and High Concentration Conditions for all Signals, and Levels (Round 2).

3.1.4 Effects of Affective State on Annoyance

An overall analysis of affective state indicated little difference in the number of highly annoyed judgments for the positive and negative affective states, as seen in Figure 3-5 (Round 1), Figure 3-6 (Round 2), and Figure 3-7 (combined data from Rounds 1 and 2).

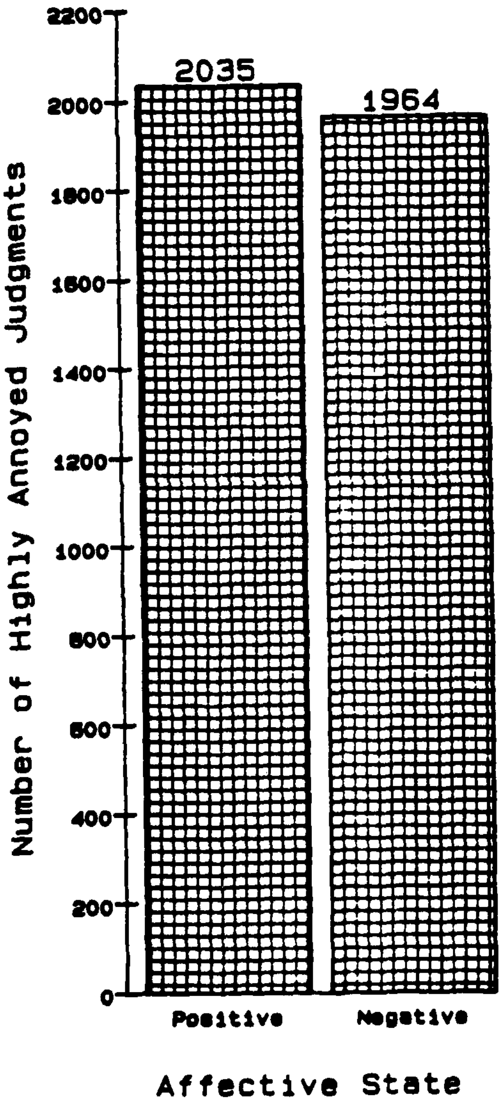


Figure 3-5: Number of Highly Annoyed Judgments for Positive and Negative Affective States for all Signals, and Levels (Round 1).

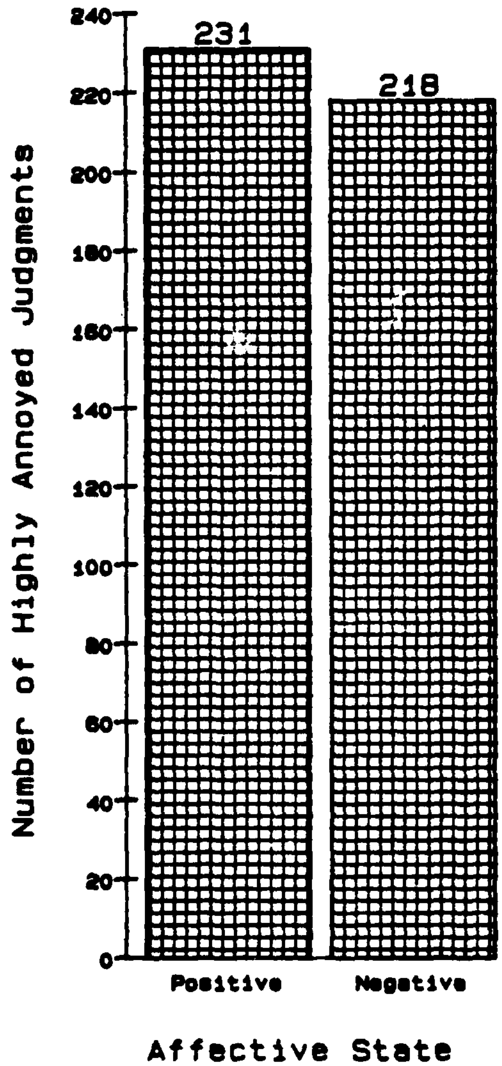


Figure 3-6: Number of Highly Annoyed Judgments for Positive and Negative Affective States for all Signals, and Levels (Round 2).

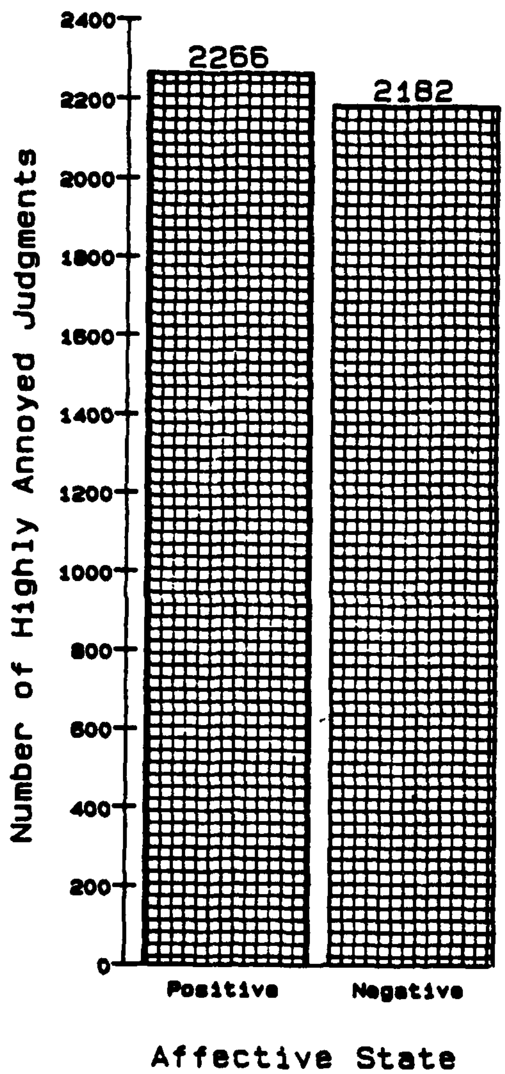


Figure 3-7: Number of Highly Annoyed Judgments for Positive and Negative Affective States for all Signals, and Levels (Rounds 1 and 2 Combined).

3.2 Inferential Analyses of Annoyance Data

The statistical significance of the findings described above was evaluated by analysis of variance. Measures taken to prepare the data for this analysis are described next.

3.2.1 Transformation of Annoyance Rating Data

Since the arbitrary representations of the categorical judgments did not lend themselves to inferential analyses, the response scale category labels were transformed into ratio scale values on an assumed unidimensional scale of annoyance via a previously performed Thurstone Case V analysis (Fidell and Teffeteller, 1981). The transformed representation of the annoyance response scale is shown in Table 3-1 and Figure 3-8.

Table 3-1: Thurstone-Scaled Representation of Annoyance Response Scale.

Transformed Representations of Categorical Judgments		
Annoyance Response Category	*Arbitrary Representation*	*Thurstone-scaled Representation*
Not at All	1	0.00
Slightly	2	2.05
Moderately	3	2.76
Very	4	3.05
Extremely	5	4.35

Even though the transformation is nonlinear, it is monotonic, so that no information is lost in converting the arbitrarily coded representations of annoyance responses to ratio scaled values in this manner. Tables 3-2 and 3-3 show the mean transformed annoyance ratings for all test conditions, collapsed across subjects and signals.

Table 3-5 show the analysis of variance results for the second round of data collection. Again, a strong effect of signal presentation level is apparent (accounting for 89% of the variance associated with signal level) and level of concentration (accounting for 33% of the variance associated with task difficulty). There was no significant interaction of signal presentation level with task difficulty in the second round of data collection.

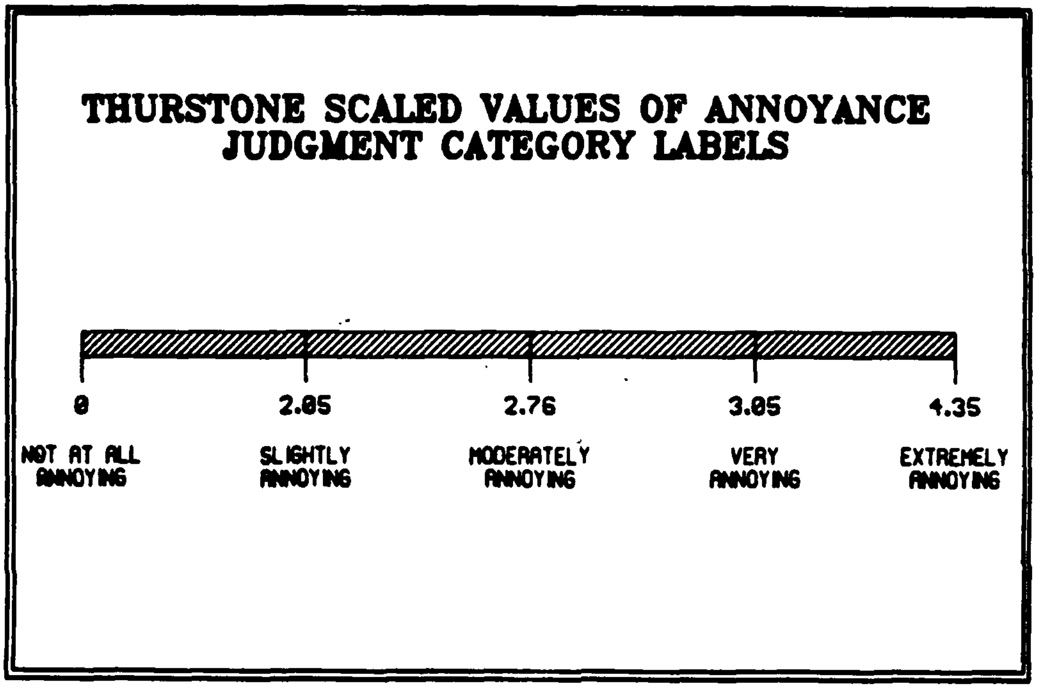


Figure 3-8: Thurstone-Scaled Unidimensional Annoyance Axis.

3.3 Predictability of Annoyance Judgments from Measures of Signal Level

Strong relationships were found between annoyance judgments and both the calculated audibility and the A-weighted levels of the test signals. Pearson product-moment correlations between annoyance judgments and audibility, and between annoyance judgments and A-weighted levels were both greater than 0.9, but there was no significant difference between them.

Table 3-2: Summary of Mean Scaled Annoyance Judgments by Test Condition (Round 1).

Cell Means Collapsed Over Signals				
Affective State	*Presentation Level*	*dB(A)*	*Concentration Levels* *Low*	*High*
Positive	1	60 dB(A)	1.093	1.269
	2	66 dB(A)	1.472	1.667
	3	72 dB(A)	1.847	2.120
	4	78 dB(A)	2.164	2.296
	5	84 dB(A)	2.538	2.532
	6	90 dB(A)	2.829	2.818
Negative	1	60 dB(A)	1.096	1.402
	2	66 dB(A)	1.468	1.716
	3	72 dB(A)	1.750	2.125
	4	78 dB(A)	2.194	2.278
	5	84 dB(A)	2.507	2.579
	6	90 dB(A)	2.821	2.830

Table 3-3: Summary of Mean Scaled Annoyance Judgments by Test Condition (Round 2).

Cell Means Collapsed Over Signals				
Affective State	*Presentation Level*	*dB(A)*	*Concentration Levels* *Low*	*High*
Positive	1	66 dB(A)	0.746	0.671
	2	72 dB(A)	1.047	0.854
	3	78 dB(A)	1.347	1.146
	4	84 dB(A)	1.621	1.420
	5	90 dB(A)	2.209	1.982
	6	96 dB(A)	3.024	2.841
Negative	1	66 dB(A)	0.775	0.781
	2	72 dB(A)	1.026	0.929
	3	78 dB(A)	1.260	1.084
	4	84 dB(A)	1.529	1.359
	5	90 dB(A)	2.138	1.789
	6	96 dB(A)	2.873	2.707

Table 3-4: Analysis of Variance of Annoyance Judgments for Round 1.

Accounting for Sources of Variance in Annoyance Judgments, Round 1					
Source of Variation	*DF*	*MS (effect)*	*MS (error)*	*F*	*Prob- ability*
Concentration	1	3.591	.303	11.86	.002
Affect	1	.057	.195	.29	---------
Level	5	40.366	.353	114.26	.000
Concentration x Affect	1	.219	.094	2.32	---------
Concentration x Level	5	.374	.041	9.10	.000
Affect x Level	5	.018	.048	.37	---------
Concentration x Affect x Level	5	.053	.050	1.06	---------
Subjects	27	.490			

Table 3-5: Analysis of Variance of Annoyance Judgments for Round 2.

Accounting for Sources of Variance in Annoyance Judgments, Round 2					
Source of Variation	*DF*	*MS (effect)*	*MS (error)*	*F*	*Prob- ability*
Concentration	1	1.889	.378	5.00	.049
Affect	1	.197	.090	2.18	---------
Level	5	26.960	.348	77.58	.000
Concentration x Affect	1	.008	.059	.13	---------
Concentration x Level	5	.074	.073	1.01	---------
Affect x Level	5	.081	.068	1.19	---------
Concentration x Affect x Level	5	.016	.058	.28	---------
Subjects	10	.803			

4. Discussion

4.1 Effects of Signal Presentation Level on Annoyance Judgments

The strong relationship observed between signal presentation level and percentage of highly annoyed judgments (as shown in Figure 3-1) lends face validity to the current observations. Figure 4-1 provides an alternative view of the relationship, plotting mean annoyance ratings (collapsed over signals, subjects, and other experimental conditions) against signal presentation levels. The linear regressions shown in Figure 4-1 account for 96% of the variance in the relationship in Round 1, and for 92% of the variance in Round 2.

Comparison of the present relationship between signal presentation level and annoyance with one derived from a prior study reveals that the annoyance of the signals presented in the current study did not grow as rapidly with presentation level as in the prior study. Figure 4-2 shows cumulative distributions of annoyance ratings collected by Fidell and Teffeteller (1981).[4] Figures 4-3 and 4-4 show comparable distributions of annoyance ratings for Rounds 1 and 2 of the present study. Note that the slopes of the cumulative distributions are considerably shallower in both rounds of data collection in the present study. In the study of Fidell and Teffeteller, the probability of judging a signal as slightly or moderately annoying doubled over a range of about 6.1 dB. In Round 1, the comparable change in signal level was about 9.7 dB; in Round 2, the change was about 15 dB. In other words, the current test subjects were less likely to be as annoyed at a given signal presentation level as those of the prior study.

One potential explanation for the greater annoyance of noise intrusions noted by Fidell and Teffeteller (1981) is that test subjects in this former experiment were engaged in a more entertaining primary task - playing video games - than in the current study. The primary (proofreading) task of the current study was considerably less entertaining, so that noise intrusions may have been less bothersome. If this explanation is accepted, it provides further (albeit indirect) support for the hypothesis that affective state influences the annoyance of noise intrusions.

---

[4]Subjects in the study of Fidell and Teffeteller (1981) used the same response scale as that of the present study, but were engaged in a different primary task and heard noise intrusions under different circumstances.

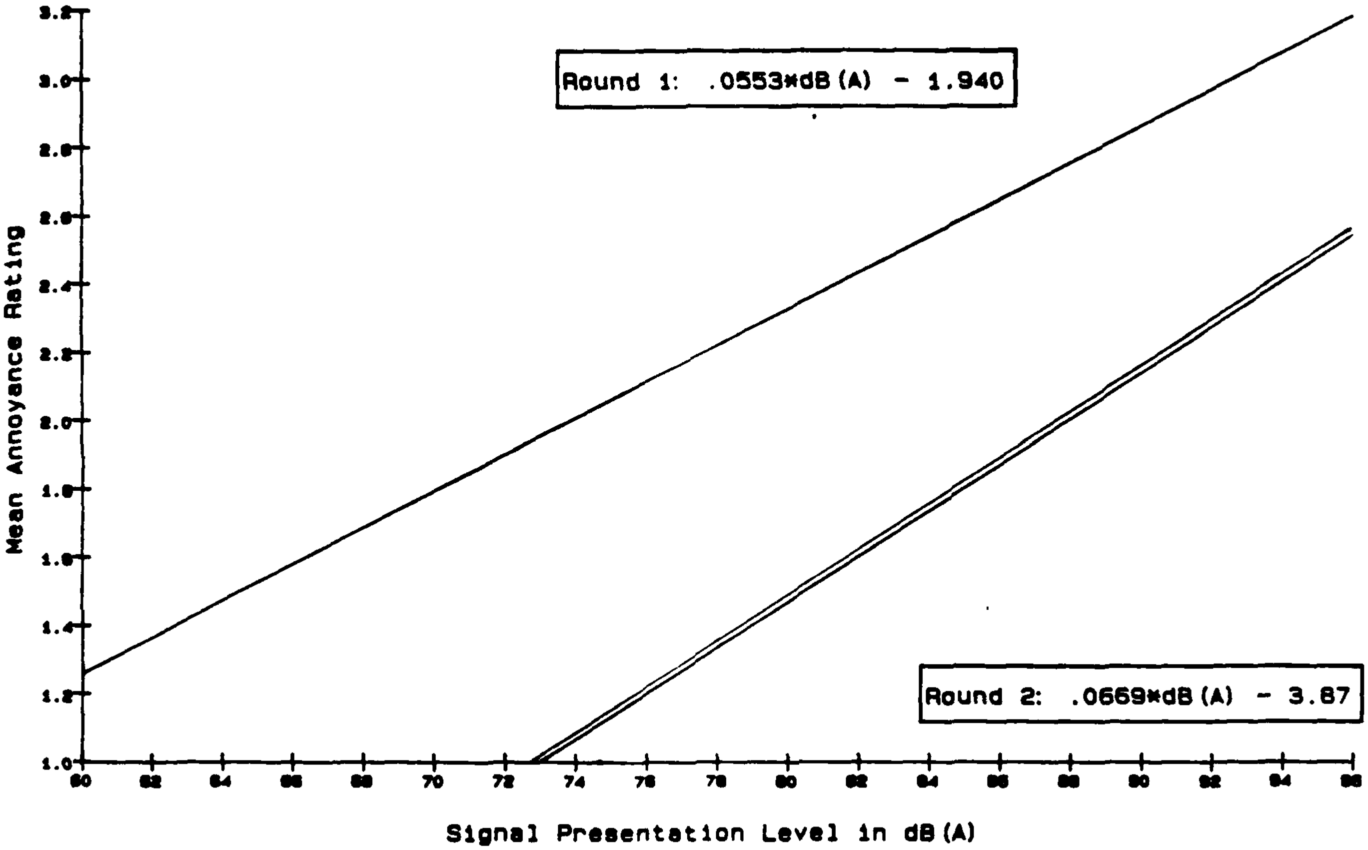


Figure 4-1: Relationship Between Mean Annoyance Ratings and Signal Presentation Levels.

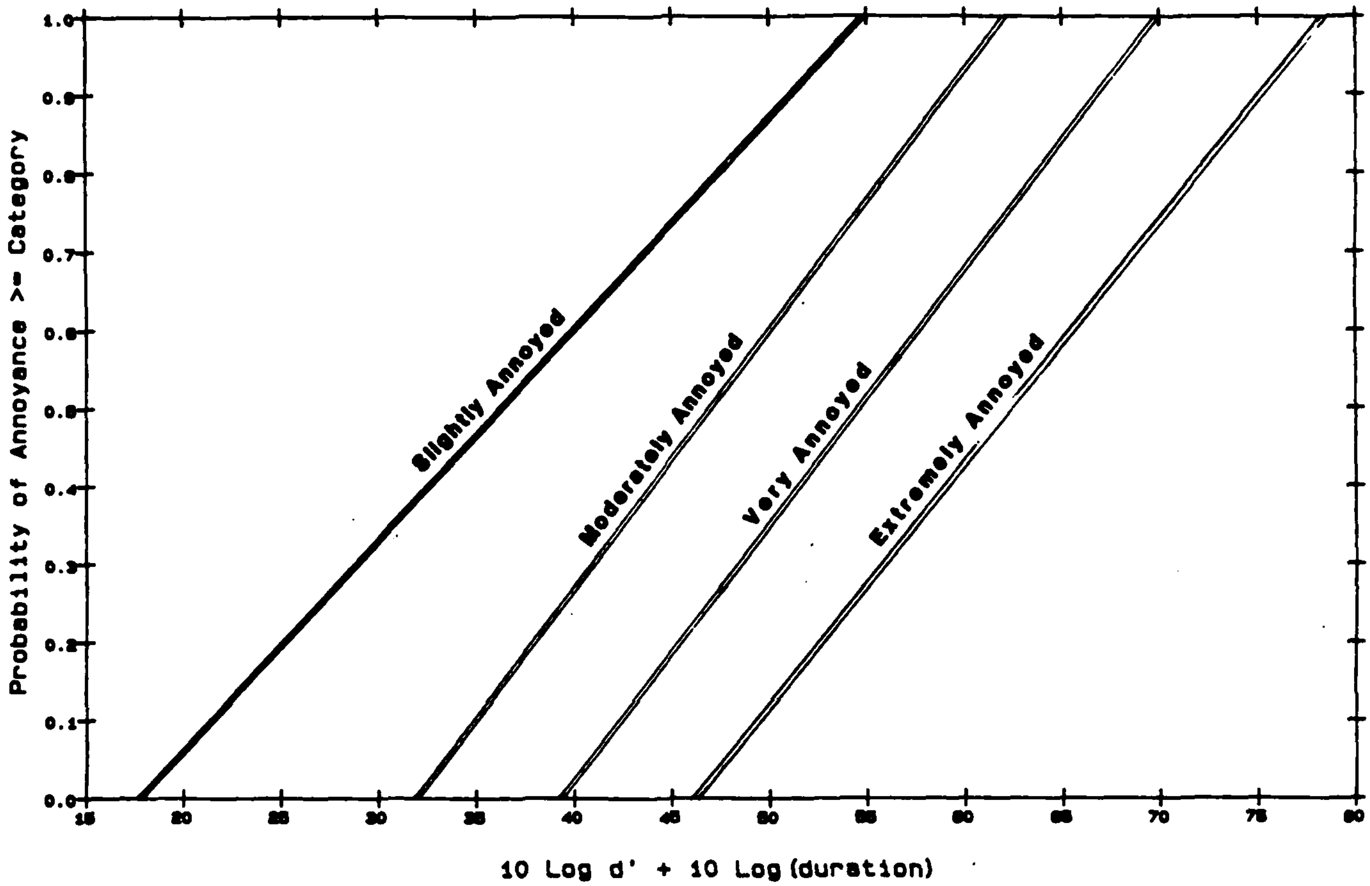


Figure 4-2: Cumulative Distributions of Annoyance Ratings Observed by Fidell and Teffeteller (1981).

4.2 Effects of Concentration on Annoyance Judgments

As noted in Section 3.1.3, Round 1 test subjects engaged in a difficult primary task were slightly (but significantly) more annoyed by noise intrusions of the same level as test subjects engaged in a less demanding primary task. However, as the difficulty of the primary task was increased in Round 2, test subjects were significantly less annoyed by noise of the same level as subjects engaged in a less demanding primary task. This observation supports the hypothesis that concentration on ongoing activities influences the apparent annoyance of noise intrusions.

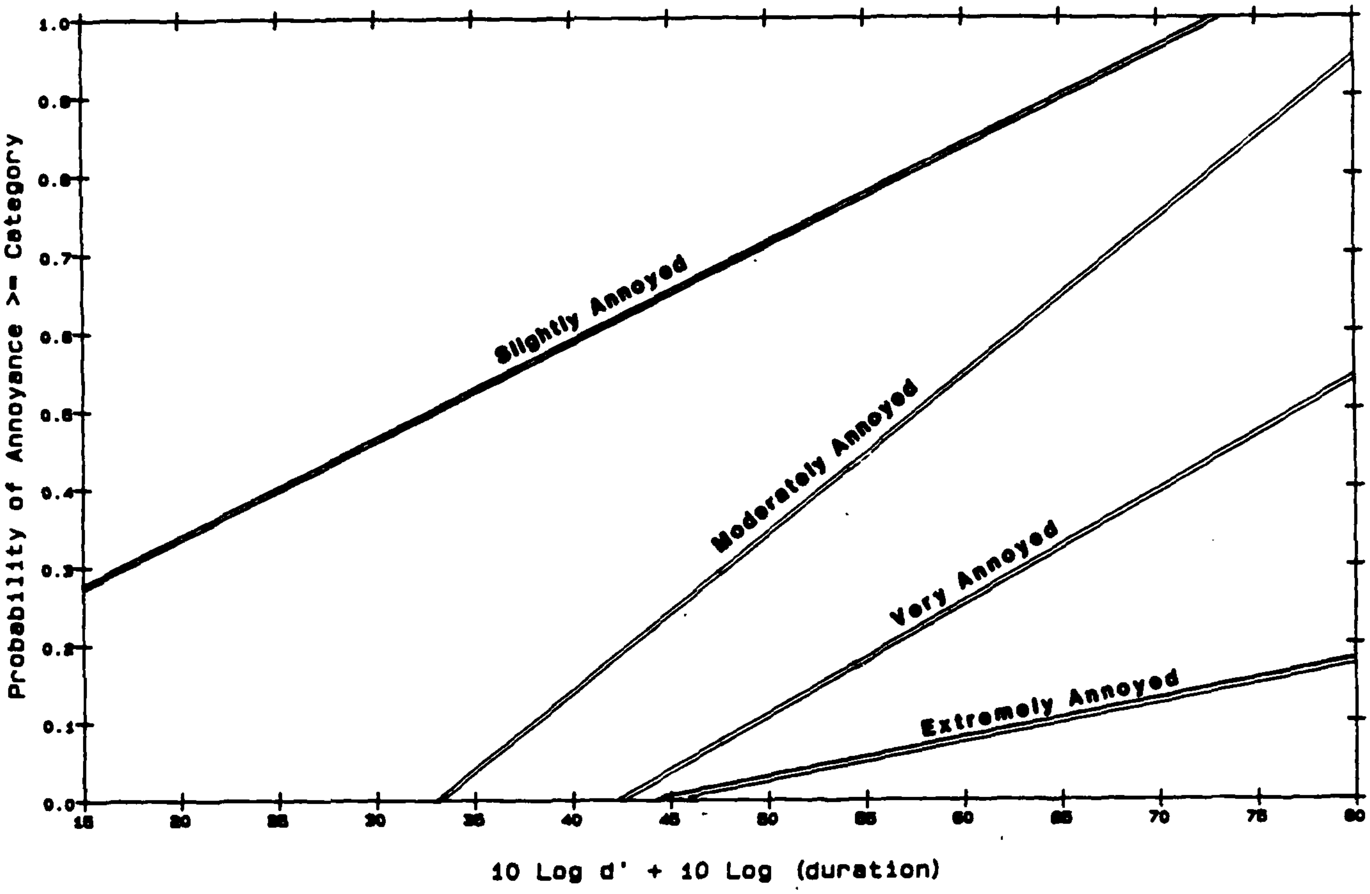


Figure 4-3: Cumulative Distributions of Annoyance Ratings Observed in Round 1.

As may be seen in Tables 3-2 and 3-3, the greatest difference in judged annoyance as a function of concentration on the primary task was somewhat less than half of a (transformed) annoyance category. The slopes of the linear regressions shown in Figure 4-1 indicate that this difference is equivalent to a difference of approximately 5 dB in signal presentation level.

These data suggest that the susceptibility of people to annoyance caused by noise intrusions may vary nonmonotonically with task difficulty, as noted in Section 1.2.2 and Figure 4-5. When people are engaged in simple tasks (i.e., those that do not tax their ability to concentrate), they may judge noise intrusions to be more annoying than when they are engaged in tasks requiring appreciable amounts of attention, simply because they are at liberty to pay attention to the noise intrusions rather than the task at hand. However, when people are engaged in complex tasks (i.e., those that do tax their ability to concentrate, as in the second round of data collection), they

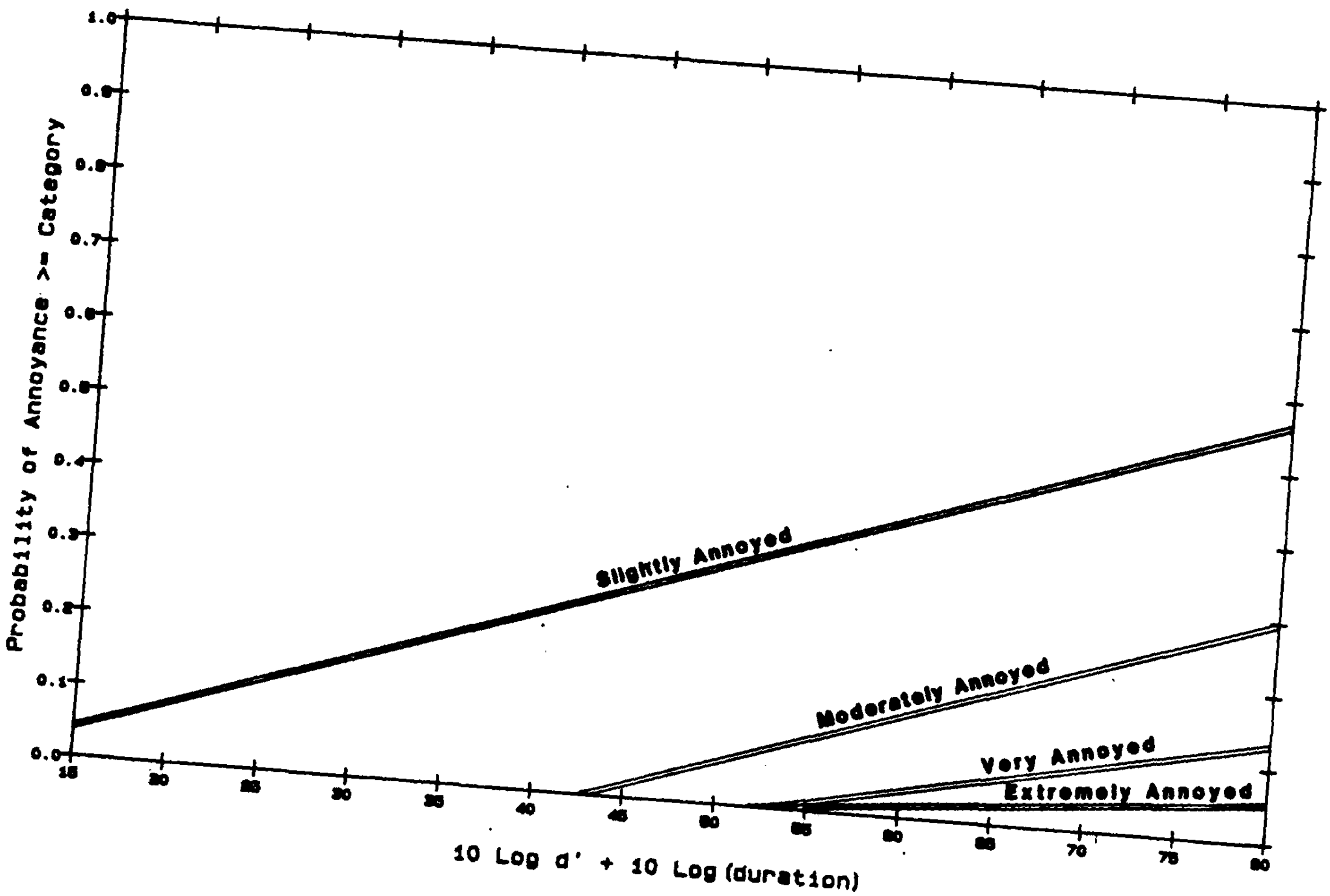


Figure 4-4: Cumulative Distributions of Annoyance Ratings Observed in Round 2.

may judge noise intrusions to be less annoying than when they are engaged in tasks that do not require appreciable amounts of attention, because they have little attentional capacity to spare for the noise intrusion. In the former case, noise intrusions may be judged annoying even though they do not affect task performance, whereas in the latter case, noise intrusions may be found annoying because they can not be permitted to affect task performance.

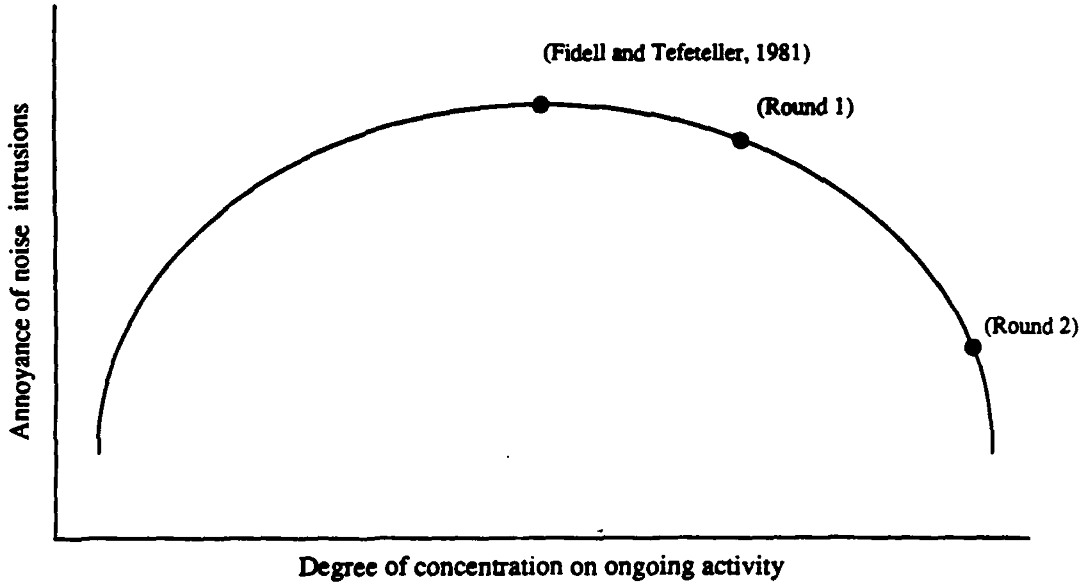


Figure 4-5: Interpretation of Effects of Task Difficulty on Annoyance of Noise Intrusions.

4.3 Ineffectiveness of Manipulation of Affect

The various measures described in Section 2.2 that were taken to manipulate subjects' affective states proved largely ineffective. Test subjects' self-ratings of mood did not change systematically to reflect these manipulations. Several post hoc analyses were conducted as described next to further test the hypothesis that the judged annoyance of noise intrusions is influenced by affective state.

4.3.1 Further Analyses of the Influences of Affective State on Annoyance Judgments

Two approaches were taken to defining test subjects' affective state independently from the planned but ineffective manipulations. In the first approach, bonus money was treated as a potential determinant of affective state. High earnings of bonus money were considered to create a positive mood, while low earnings of bonus money were considered to create a negative mood. In the second approach, the subjects' self-ratings of mood were used as indicators of positive and negative affective states.

4.3.2 Bonus Earnings as a Determinant of Annoyance Judgments

Subjects were ranked by the total amount of bonus money earned during the experiment. The annoyance judgments of the quarter of the subjects earning the most money and the quarter of the subjects earning the least amount of money were then contrasted. Mean annoyance judgments were calculated for each group at each presentation level as seen in Tables 4-1. T-tests of the mean differences revealed a significant difference in the average annoyance judgments when categorizing subjects by earnings ($t_{5df} = -10.78$, $p = .001$).

Table 4-1: Mean Annoyance Ratings for Subjects Earning High vs Low Bonus Money at Each Signal Presentation Level.

Mean Annoyance Ratings		
Presentation Level	*High Earnings*	*Low Earnings*
1	.929	1.764
2	1.285	1.943
3	1.633	2.307
4	1.983	2.569
5	2.387	2.881
6	2.845	3.291
Mean	1.844	2.459
Standard Deviation	.709	.575

4.3.3 Self-Rated Mood as a Determinant of Annoyance Judgments

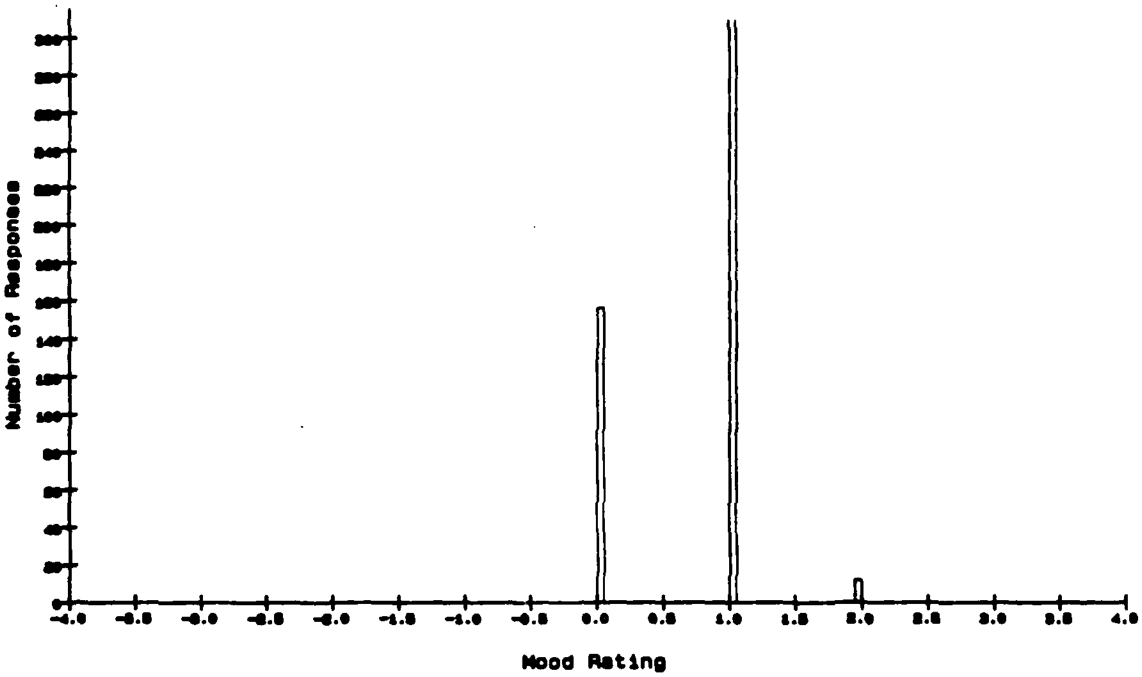
The values of the mood rating scale on which subjects rated themselves at the end of every block of trials ranged from -4 (extremely bad mood) to +4 (extremely good mood). Frequency histograms were developed for all subjects' usage of this scale. Samples of these histograms may be seen in Figure 4-6. Examination of these histograms revealed that 12 of the 28 subjects did not use many of the scale categories, but instead clustered their mood ratings around neutral or at one of the two extremes (positive or negative). These subjects' data were, therefore,

omitted from this analysis, on the grounds that their mood reports did not vary appreciably during the course of the experiment.

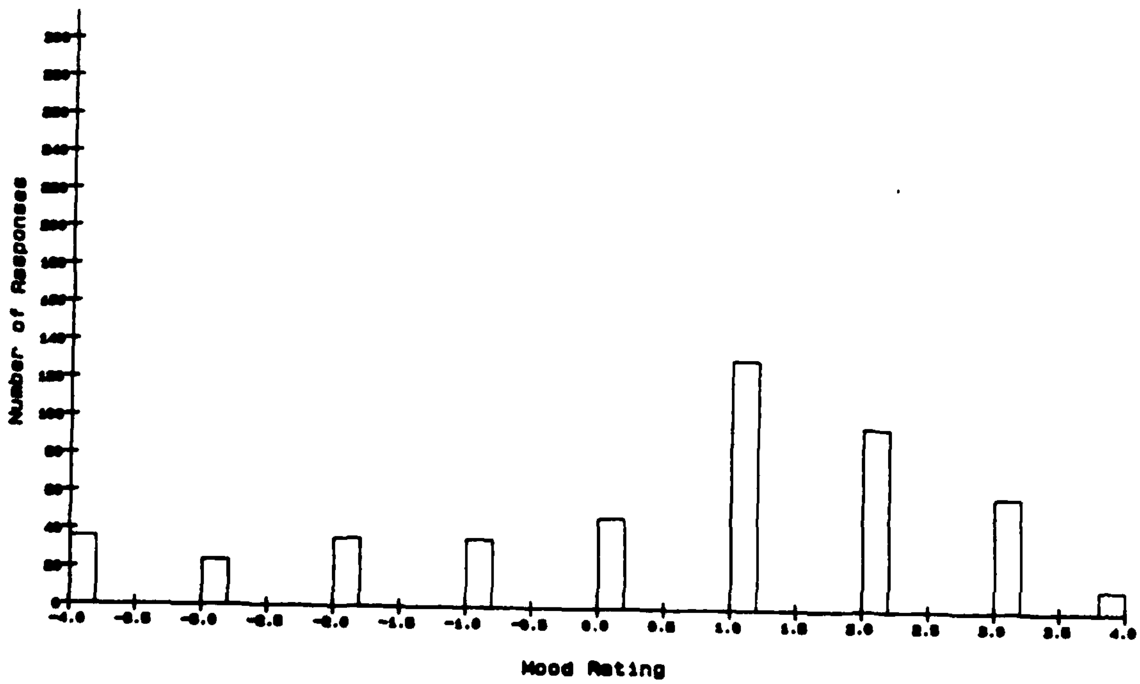
Mean annoyance ratings were calculated for the remaining 16 subjects' "good mood" and "bad mood" trial blocks. "Good mood" blocks of trials were defined as those with mood ratings greater than or equal to +2.0. "Bad mood" blocks of trials were defined as those with mood ratings less than or equal to -2.0. Trial blocks with mood ratings intermediate between +2 and -2 were ignored in this analysis. Table 4-2 summarizes the mean annoyance ratings for each subject's trials within the identified "Good Mood" and "Bad Mood" blocks, collapsed over signals, presentation levels, and experimental conditions.

A comparison of the annoyance ratings of the two groups of trial blocks revealed that annoyance judgments made in "Good Mood" blocks (mean = 2.39) were significantly lower than annoyance judgments in the "Bad Mood" blocks (mean = 2.76) ($t_{15df} = -3.22$, $p < .05$).

This finding suggests that annoyance of acoustic signals is considered greater when people are in a negative mood, thus supporting the hypothesis that affective state influences judgments of annoyance.



Example of a subject with neutral self-ratings of mood.



Example of a subject with extreme self-ratings of mood.

Figure 4-6: Sample Histograms of Use of Mood Rating Scale.

Table 4-2: Mean Annoyance Ratings on Trials With Extreme Mood Ratings.

Mean Annoyance Ratings		
Subject No.	*"Good Mood" Trials*	*"Bad Mood" Trials*
6	3.18	3.37
8	2.88	3.56
11	1.68	3.32
12	2.45	2.41
13	2.32	2.68
16	2.56	2.81
17	2.44	2.87
18	2.98	3.04
19	1.67	2.26
21	2.17	2.63
22	1.35	1.44
27	2.55	2.55
28	2.18	3.32
29	1.98	1.97
31	3.50	3.84
32	2.28	2.14
Mean	2.39	2.76 **
Standard Deviation	.57	.64

**p <.05

5. Conclusions and Recommendations

This experiment produced evidence supporting the reasonableness of the hypotheses under test. Even though the evidence was consistent with the hypotheses, it was not strong enough to be considered conclusive. Although stronger confirmation of the model's assumptions could probably be derived from additional laboratory testing (for example, in a between subjects experimental design, or with a more entertaining primary task), it is more useful for present purposes to seek additional evidence in field settings. The sort of field tests that could produce such evidence would require simultaneous, real time monitoring of noise exposure and annoyance. Immediate annoyance judgments would be accompanied by longer term measures of affective state and activity at the time of noise intrusions. Miniaturized, computer-based apparatus would have to be developed to record such information.

The analyses presented in this report focused narrowly on issues related to the major hypotheses under test; they do not exhaust the information produced by the current experimentation. Of particular interest are the unanalyzed relationships between performance in the primary task and signal presentation level, affect, and task difficulty. Other unanalyzed issues that warrant further exploration include information about the prediction of signal-specific annoyance, and the generalization of annoyance of individual noise intrusions to the annoyance of long term noise exposure.

The results of the present and any additional analyses that may be performed on the current data set should be incorporated into a software module suitable for inclusion in ASAN (cf. Fidell, Reddingius, Harris, Kugler, 1989) for prediction of the annoyance of individual noise intrusions.

References

Fidell, S. and Green, D., "A Rationale and Plan for Developing Improved Means of Predicting Aircraft Noise Annoyance," BBN Report 6751, October, 1988.

Fidell, S., Green, D., Schultz, T., and Pearsons, K., "A Strategy for Understanding Noise-Induced Annoyance," BBN Report 6337, 1988.

Fidell, S., Horonjeff, R., Teffeteller, S., and Green, D., "Effective masking bandwidths at low frequencies," J. Acoust. Soc. Am., 73(2): 628-638 (1982).

Fidell, S., Horonjeff, R., Mills, J., Baldwin, E., Teffeteller, S., and Pearsons, K., "Aircraft noise annoyance at three joint air carrier and general aviation airports," J. Acoust. Soc. Am., 77(3): 1054-1068 (1985).

Fidell, S., Pearsons, K., Grignetti, M., and Green, D., "The Noisiness of Impulsive Signals," J. Acoust. Soc. Am., 48(6) 1304-1310 (1970).

Fidell, S., Reddingius, N., Harris, M., and Kugler, B.A., "Executive Summary of Initial Development of an Assessment System for Aircraft Noise (ASAN)," BBN Report 6624, 1989.

Fidell, S., and Teffeteller, S., "Scaling the Annoyance of Intrusive Sounds," J. of Sound and Vib., 78(2): 291-298 (1981).

Fidell, S., Teffeteller, S., Horonjeff, R., and Green, D., "Predicting Annoyance from Detectability of Low Level Sounds," J. Acoust. Soc. of Amer., 66(5): 1427-1434 (1979).

Green, D., and Swets, J., Signal Detection Theory and Psychophysics, New York: John Wiley and Sons, 1966.

Horonjeff, R., Fidell, S., Teffeteller, S., and Green, D., "Behavioral awakening as functions of duration and detectability of noise intrusions in the home," J. Sound and Vib., 84(3): 327-336 (1982).

Swets, J.A., Signal Detection and Recognition by Human Observers, John Wiley and Sons, Inc., New York, 1964.

Tabachnick, B., and Fidell, L., "Using Multivariate Statistics," 2nd ed., Harper and Row, New York, 1989.

Tanner, W.P., and Birdsall, T.G., "Definitions of d' and η as Psychophysical Measures," J. Acoust. Soc. Am., 30(10): 922-928 (1958).

Appendix A
Instructions to Test Subjects

The following instructions were displayed on the test subjects' terminals when they were first seated in the anechoic chamber:

You are about to take part in a study in which you will be asked to count the number of times that certain characters appear on the screen on which you are reading these instructions. You will also be asked on a number of occasions to tell us your opinion about how annoying you find certain sounds.

Press return now to see the detailed instructions for these two tasks.

(RETURN PRESSED)

DETAILED INSTRUCTIONS FOR THE FIRST (CHARACTER SEARCH) TASK:

This task will be divided into groups of separate trials. At the start of each trial you will see at the top of the screen the characters that you should look for. Press return to see what this announcement will look like.

(RETURN PRESSED)

```
Characters to look for:  C  A
```

Next, you will see several lines of characters. Press return again for an example.

(RETURN PRESSED)

```
A  W  T  A  J  A  J  V
K  W  M  E  W  A  A  E
M  M  W  W  A  W  A  J
W  M  E  J  C  A  J  F
E  L  M  L  F  A  W  A
J  L  J  W  M  M  W  J
F  C  W  F  C  P  M  D
Z  Q  C  A  U  P  I  S
```

After a brief period of time, you will be asked to enter the number of times that you saw the characters you were supposed to be looking for. Keep in mind that you are to search for each character separately from the other target characters. Make sure that you look for individual occurrences of the target characters, not just for a particular sequence of target characters.

Press return to see how you will be asked to enter your answers.

(RETURN PRESSED)

Number of appearances of C: ___

The computer will wait only a short time for your first answer (the number of times you saw the first specified character). You will then be asked to supply the number of any other target characters that you may have been asked to look for. Again, you will have only a short time to type the number of times you saw any other target character that you were asked to look for. Press return to continue.

(RETURN PRESSED)

Sometimes the character search task will be easy, and sometimes it will be difficult. However, if you can provide the correct count of the number of times that each of the characters you were looking for actually appeared on the screen during the trial, you will be awarded a bonus in addition to your hourly pay of $5.00 per hour. Correct answers are worth either two cents or four cents for individual target characters in different experimental conditions. Press return to continue.

(RETURN PRESSED)

The value of correct answers also increases with the difficulty of the proofreading task. If you correctly determine the number of times that ***both*** target characters appear on a trial when you are asked to look for two targets, then the bonus for the second correct answer is doubled. If you correctly determine the number of times that ***two of three*** target characters appear on a trial when you are asked to look for three targets, then the bonus for the second correct answer is doubled. If you answer the third correctly, the bonus for that answer is tripled.

On certain trials, called "triple bonus trials," the sum of the bonuses you earn for the trial is tripled.

These bonuses add up quickly, so that over the course of the entire experiment, you can more than double your hourly rate of pay if you concentrate on what you're doing. Press return to continue.

(RETURN PRESSED)

There is no penalty for guessing, so feel free to provide an answer even if you're not certain about it. The only penalty is for not answering at all. Each time you fail to provide any answer, a fine of twenty five cents will be subtracted from your accumulated bonus. Bonuses and fines do not affect your hourly pay in any way.

The computer will keep track of your bonus payment, which will be paid to you at the end of each day's testing.

Press return to continue.

(RETURN PRESSED)

After the experiment begins, you will be given breaks at least twice per hour, and you can stop the experiment at any time by simply pushing the button on the desk next to the computer. All activity for both you and the other subject seated in the lab with you will be stopped. The experimenter can also hear you and will be watching you through the TV camera in front of you. If all else fails, you can simply walk out of this chamber. Unless there is some good reason for doing so, however, we would prefer that you wait until a set of trials ends before standing up so as not to disturb others.

Press return when you are ready to see the instructions for the second job that you will be expected to perform.

(RETURN PRESSED)

DETAILED INSTRUCTIONS FOR THE SECOND (ANNOYANCE RATING) TASK:

Sometimes while you are in the midst of the character search task you will hear a sound of varying loudness, such as those that you will hear after you finish reading these instructions.

Keep working on the character search task when such sounds occur. After you have typed your answers in the usual manner, you will be asked to rate the annoyance of the sound that occurred while you were working on the character search task. In other words, you will be asked to judge how annoyed you were by hearing the sound while you were working on the character search task. Press return for an example of the annoyance rating you will be asked for.

(RETURN PRESSED)

The prompt for your annoyance rating will look like the following:

```
Annoyance rating:  1=Not At All Annoying
                   2=Slightly Annoying
                   3=Moderately Annoying
                   4=Very Annoying
                   5=Extremely Annoying

Your annoyance rating (1-5) of the sound: ___
```

Press return for your final instructions.

(RETURN PRESSED)

At the end of a block of trials, about 15 minutes long, you will be asked for two more pieces of information. The data will be a rating of your mood, and of the difficulty of the proofreading task during the preceding block of trials. Neither of these ratings affects your bonus or the tasks you will be given later, so please rate these as accurately as possible. Failure to provide an answer will result in a fine of twenty-five cents to be deducted from your bonus. Press return for an example of the first rating you will be asked for.

(RETURN PRESSED)

The prompt for your rating of your mood will look like this:

```
Mood rating:    1=Extremely Bad
                2=Very Bad
                3=Moderately Bad
                4=Slightly Bad
                5=Neutral
                6=Slightly Good
                7=Moderately Good
                8=Very Good
                9=Extremely Good

Your rating (1-9) of your current mood: ___
```

Press return for an example of the second rating you will be asked for.

(RETURN PRESSED)

The prompt for your rating of the difficulty of the block of trials will look like this:

```
Difficulty rating:  1=Very easy
                    2=Fairly easy
                    3=Average difficulty
                    4=Fairly difficult
                    5=Very difficult

Your rating (1-5) of the difficulty of the character search task: __
```

Press return for your final instructions.

(RETURN PRESSED)

Most of the time that you participate in this experiment there will be another test subject in the anechoic chamber with you. Please do not talk to the other test subject while the test is in progress, since this could disturb both your concentration and that of the other test subject. The other test subject's screen will not show the same information as yours, so there is nothing to be gained from communication.

Press return again for a summary of your instructions.

(RETURN PRESSED)

You have two main jobs to perform:

1) Search for target characters mixed with others; and

2) Rate the annoyance of sounds heard while performing the character search.

Although there will usually be enough time for you to type your answers, you should provide them as soon as the computer asks for them, because there may not be time for the computer to register your answers before the next trial starts.

Bonuses are paid for correct answers, but there are no penalties for incorrect answers, so it's worth your while to guess even if you're not sure of an answer. Fines are deducted from your total bonus for not answering at all.

If you have any questions, now is the time to ask the experimenter.

Appendix B
Procedures for Recruitment and Treatment of Test Subjects

Advertisements for test subjects were placed with local college employment offices and in a local circulation newspaper. Prospective subjects were offered an hourly rate of $5.00 to participate in the study, payable weekly or upon termination of participation. In addition, they were eligible for performance bonuses as described in their instructions, payable daily.

All test participants were audiometrically screened for hearing levels within 10 dB of audiometric zero by means of von Bekesy audiometry. Audiograms also were collected by the same means at the completion of participation in the study. The maximum absolute sound pressure level to which test subjects were exposed in the course of the study was hardware limited to 90 dB(A) for a single event duration no greater than ten seconds. Test equipment was acoustically calibrated at the beginning of each data collection session.

Data were collected over a period of two weeks in daily sessions that lasted approximately 2 hr each. A minimum of two 5-min rest breaks per hour were provided during which test subjects were required to leave the anechoic chamber.

Informed consent was obtained through standard means, including

1) provision of a hard copy description of the nature of the tests to be conducted and the test protocol;

2) an opportunity to ask questions and discuss the nature of the experimentation;

3) written reassurances a) that test subjects were free to terminate their participation in the study without prejudice at any time during the course of experimentation; b) that the sounds to which they were exposed did not constitute any meaningful risk of hearing damage; c) that no medical treatment was to be provided in conjunction with participation in the study; d) that subjects would receive no benefit of participation beyond their remuneration; and e) that the results of the experimentation would eventually be made available as a public document.

Appendix C
Description of Software for Test Administration

Two sets of programs were needed to administer the current experiment: one set of programs that actually controlled test administration in real-time, and one set of utility programs that produced input for the real time programs. As described in Section 2.2, the combinations of conditions administered in the 1000 trials of the experiment were sufficiently complex that a program was needed to generate parameter settings. This software produced 40 files, each containing 25 sets of parameter values which fully described all conditions for a trial.

Two other utility programs were also needed to prepare for data collection: an analog-to-digital signal recording program and an audibility calculation program. The former program captured an analog signal reproduced on a tape recorder. The output of this program was a file containing digitized data that could be played back through a 12-bit D/A converter to reproduce the original signal. The latter program calculated both A-weighted sound pressure levels and d' values for each signal. A-level information was used to compute attenuation levels at which signals were reproduced during the course of the experiment. Values of d' were used to assist in signal selection. Both of these programs were created through minor revisions of existing software.

The real-time software for administering experimental conditions controlled the hardware shown in Figure C-1. This software was written in five sections intended for concurrent execution: (1) an experimenter process; (2) process instructions; (3) and (4) the two subjects' processes; and (5) a signal generation process. The experimenter process was the controlling (parent) process that started all other processes and directed them to perform specific tasks. This main process took its input from command files generated by the utility software described above. These command files provided parameter values that the experimenter process passed along to its subprocesses via the process instructions section.

The two subjects' processes, which controlled all input and output from the test subjects' terminals, were essentially identical. They were coded as separate processes since they administered different experimental conditions while running, and waited for different start signals from the parent process. These processes took parameter values from the process instructions section to determine the type of matrix to display on the subject's terminal (i.e., the number of target characters, the size of the matrix, the similarity of the target and background characters to each other, the duration of the display of the matrix, and the number of times the target characters appeared in the matrix).

The subjects' processes randomly created display matrices with specified characteristics for

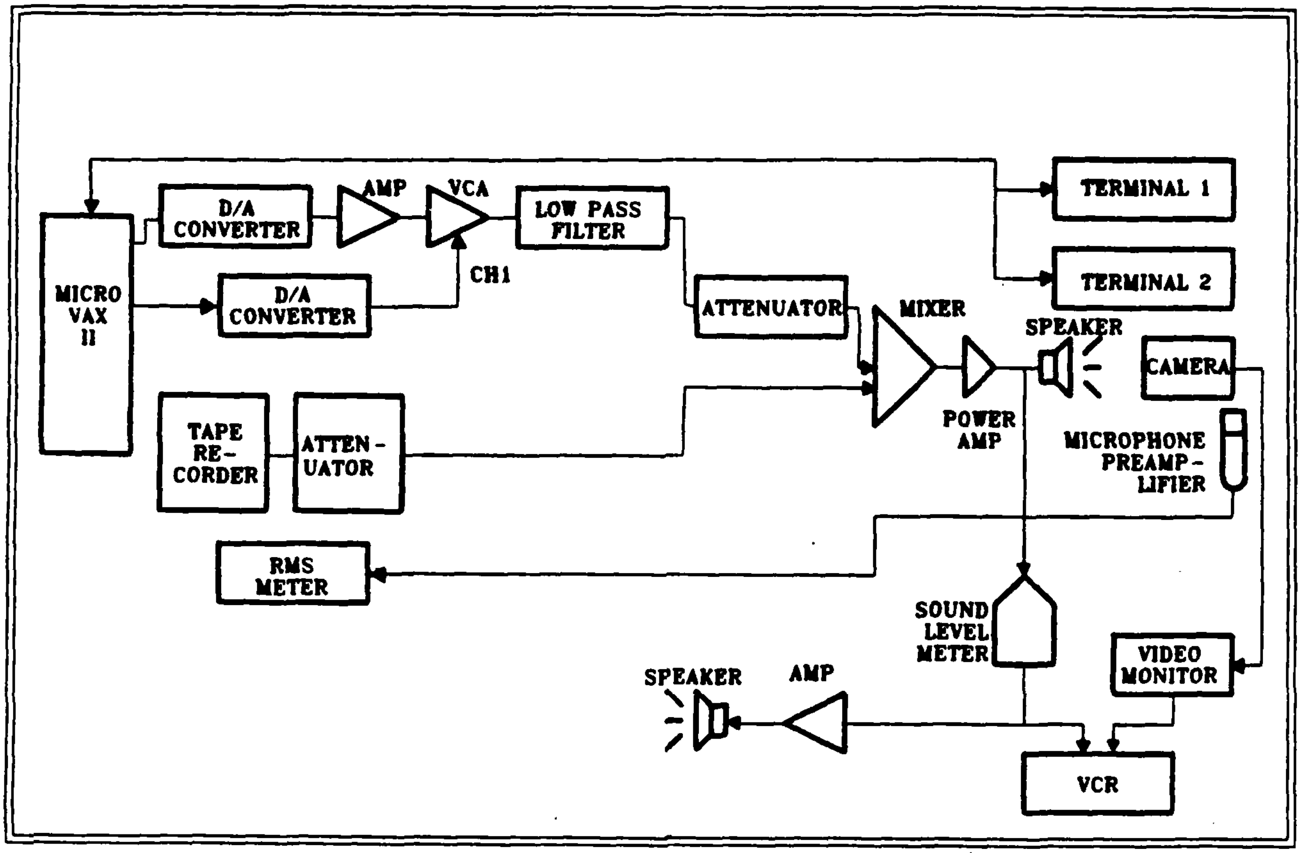


Figure C-1: Hardware controlled by real time software to reproduce test signals.

each trial. The process instructions indicated the bonus to be awarded for correct answers, the presence or absence of a response delay before the subject was prompted for an answer, and presence or absence of lock out and triple bonus conditions. All of these parameters affected the way that the subjects' processes asked for responses to the character search task and the bonuses awarded for these answers. The process instructions also indicated whether a signal was to be played on a particular trial, and whether a trial was the last trial of the block. The latter information was needed to permit the subjects' processes to elicit annoyance, mood and difficulty ratings at the proper times.

When the subjects' processes finished all these tasks, they signaled the controlling processes and in turn awaited a signal to start the next trial. The two subjects' processes could finish at different times (since they were executing different sets of conditions), but the experimenter's process kept them synchronized by waiting 30 s (3 s longer than the longest trial) before starting both of the processes for the next trial.

The other real-time program was the signal generation process. This process also took its input from the process instructions. The only parameters that the signal generation processes needed were the number of the signal to be played and the level at which it was to be played. The number of the signal corresponded to the name of a file created by the analog-to-digital (A/D) utility program containing the digitized. The selected file was read through the digitial-to-analog (D/A) converter to a voltage controlled amplifier (VCA) that attenuated it to its proper level. The level corresponded to the VCA control voltage required to reproduce a particular signal at a specified level. This control voltage was sent to the VCA via an IEEE-488 bus. The signal generation process was awakened only on signal presentation trials.

The last program used in the administration of the experiment was not controlled by the main process. This was the program that displayed the instructions for the experiment on the subjects' terminal screens. This program allowed subjects to read the instructions at separate paces, and alerted the experimenter when both the subjects were done.

Appendix D
Acoustic Characteristics of Signals Presented to Subjects for Annoyance Judgments

This Appendix contains graphic representations of the temporal and spectral characteristics of the ten test signals described in Table D-1:

Table D-1: Description of Test Signals.

Signal	*Description*
1	Aircraft flyover (Boeing 720-B recorded outdoors)
2	Aircraft flyover (DC-8 landing recorded indoors)
3	Aircraft flyover (Boeing 720 recorded outdoors)
4	Aircraft flyover (Boeing 707 landing recorded outdoors)
5	Aircraft flyover (Boeing 727 landing recorded outdoors)
6	Aircraft flyover (DC-8 landing recorded outdoors)
7	Motorcycle driveby (recorded indoors)
8	Truck passby (recorded outdoors)
9	One kHz octave band of noise
10	Five kHz one-third octave band of noise

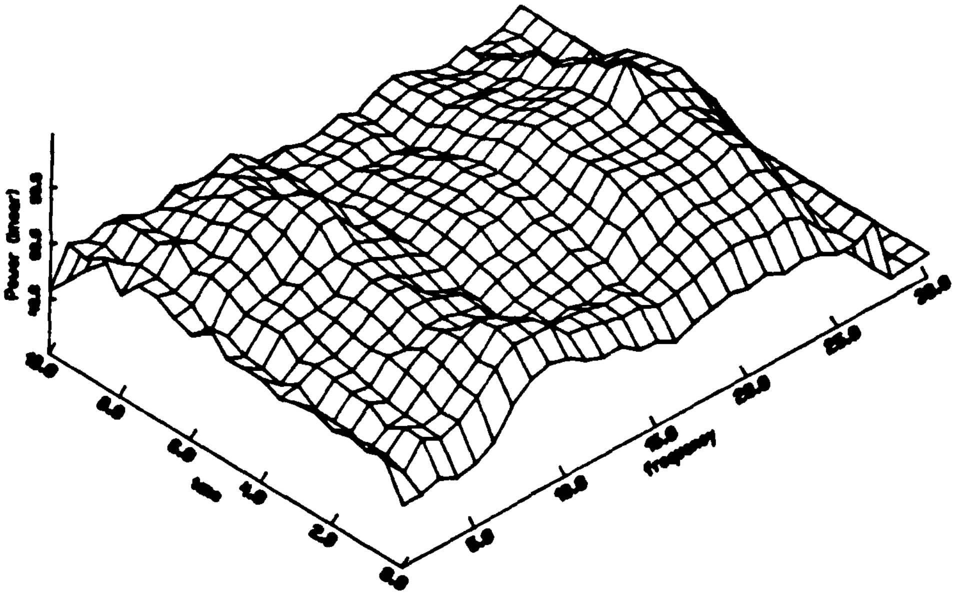


Figure D-1: Isometric View of One-Third Octave Band Spectrum of Signal 1 at Half-Second Intervals.

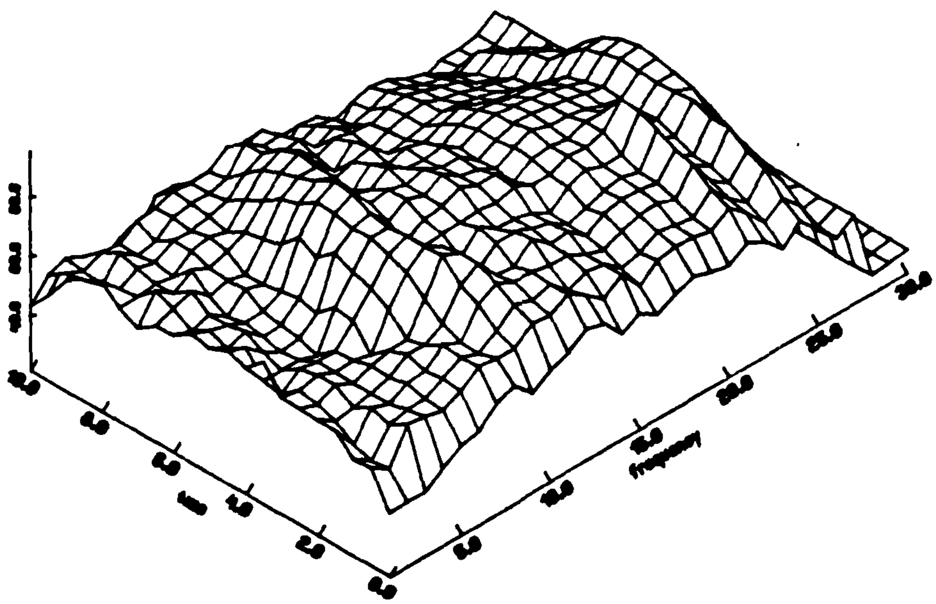


Figure D-2: Isometric View of One-Third Octave Band Spectrum of Signal 2 at Half-Second Intervals.

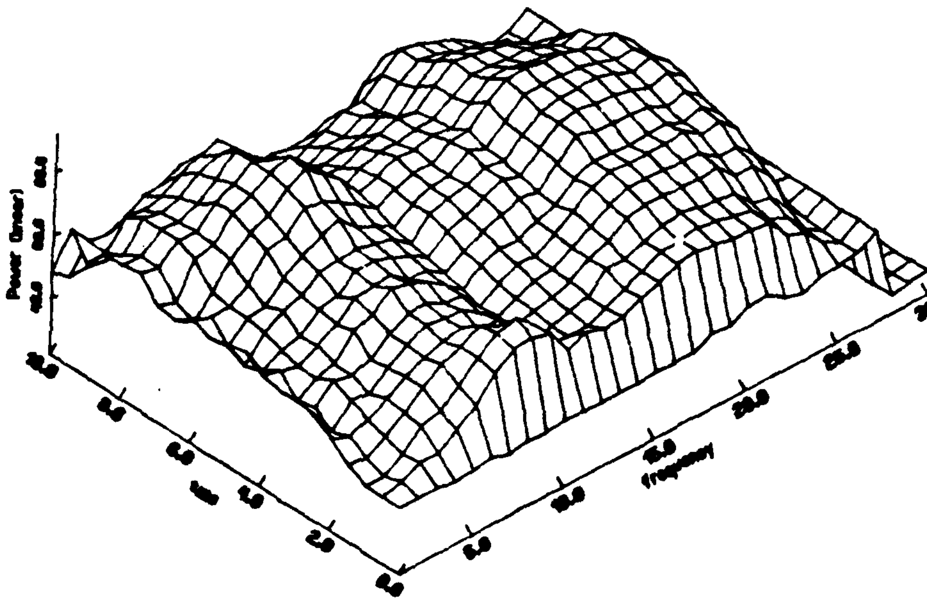


Figure D-3: Isometric View of One-Third Octave Band Spectrum of Signal 3 at Half-Second Intervals.

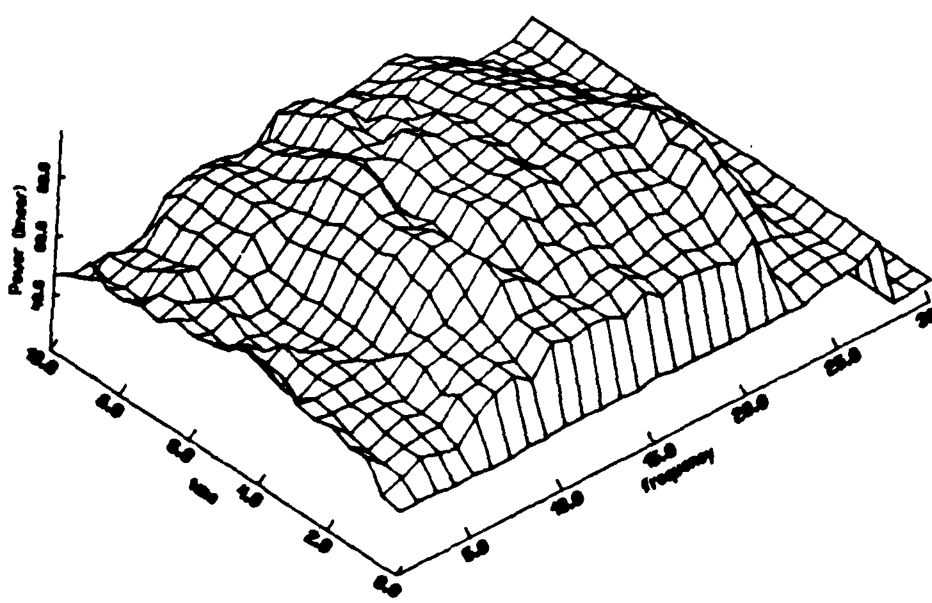


Figure D-4: Isometric View of One-Third Octave Band Spectrum of Signal 4 at Half-Second Intervals.

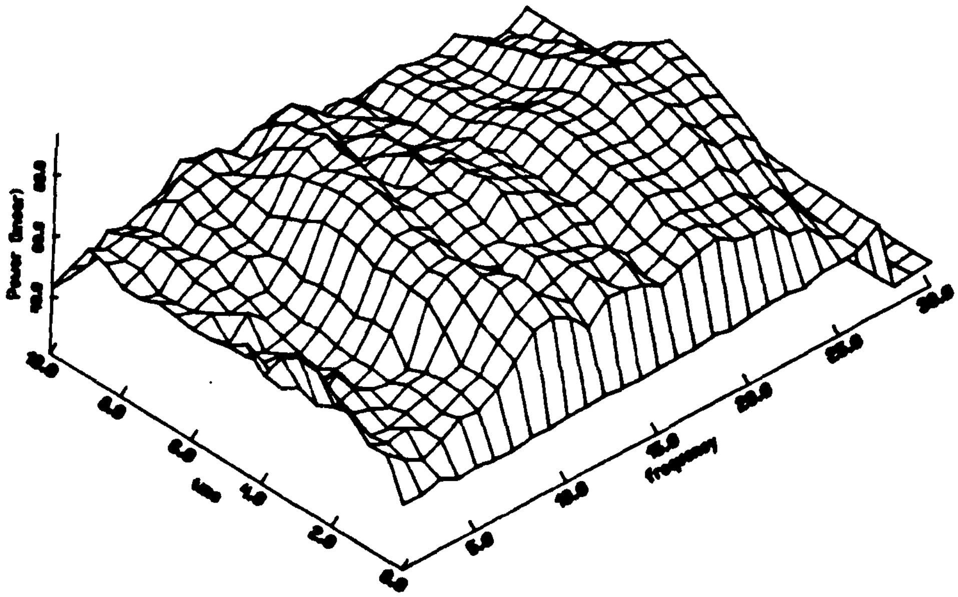


Figure D-5: Isometric View of One-Third Octave Band Spectrum of Signal 5 at Half-Second Intervals.

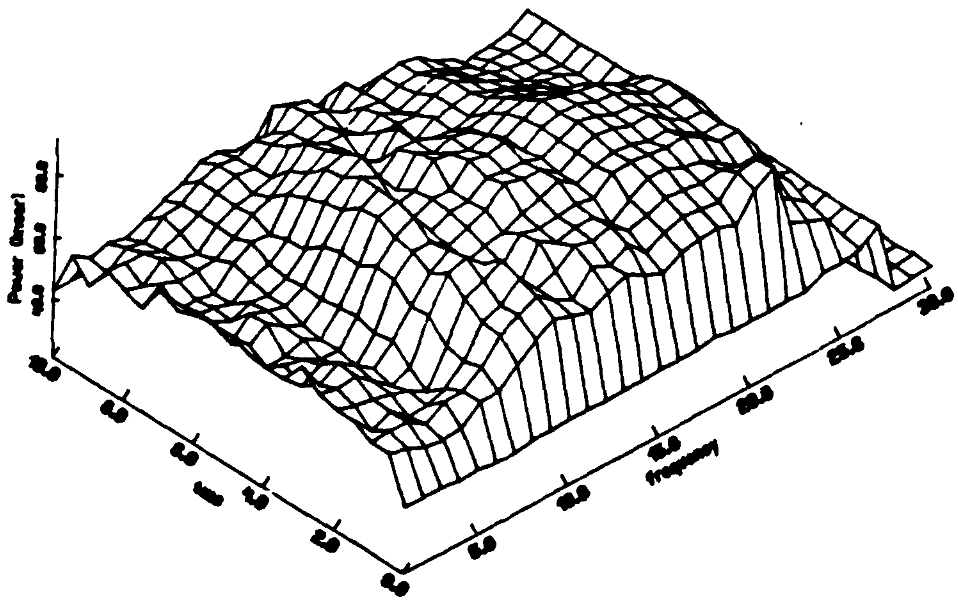


Figure D-6: Isometric View of One-Third Octave Band Spectrum of Signal 6 at Half-Second Intervals.

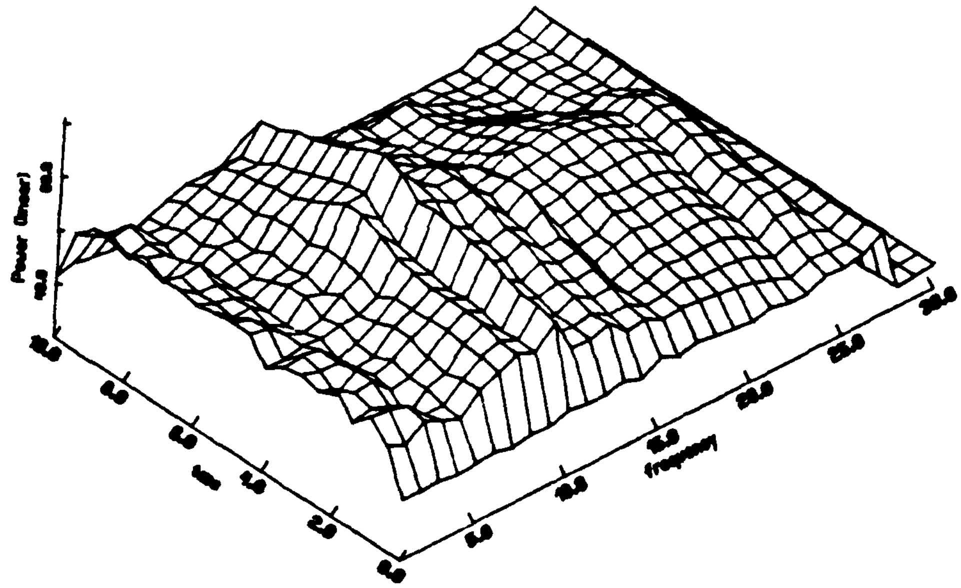


Figure D-7: Isometric View of One-Third Octave Band Spectrum of Signal 7 at Half-Second Intervals.

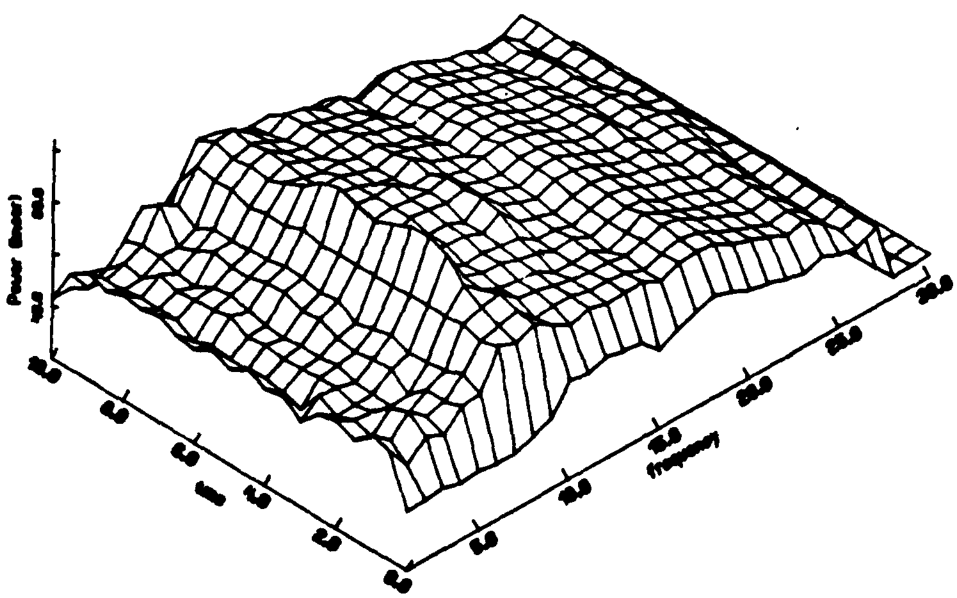


Figure D-8: Isometric View of One-Third Octave Band Spectrum of Signal 8 at Half-Second Intervals.

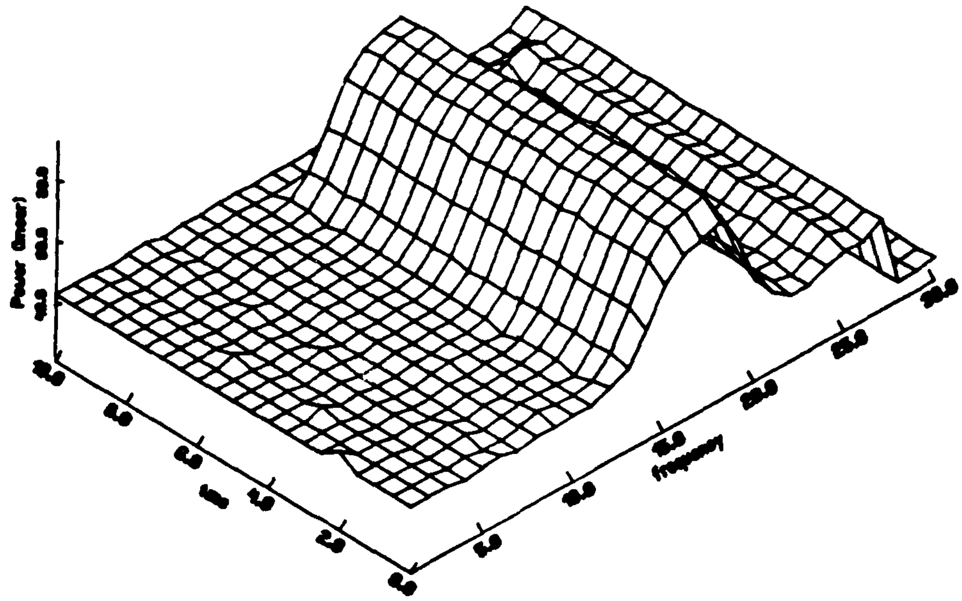


Figure D-9: Isometric View of One-Third Octave Band Spectrum of Signal 9 at Half-Second Intervals.

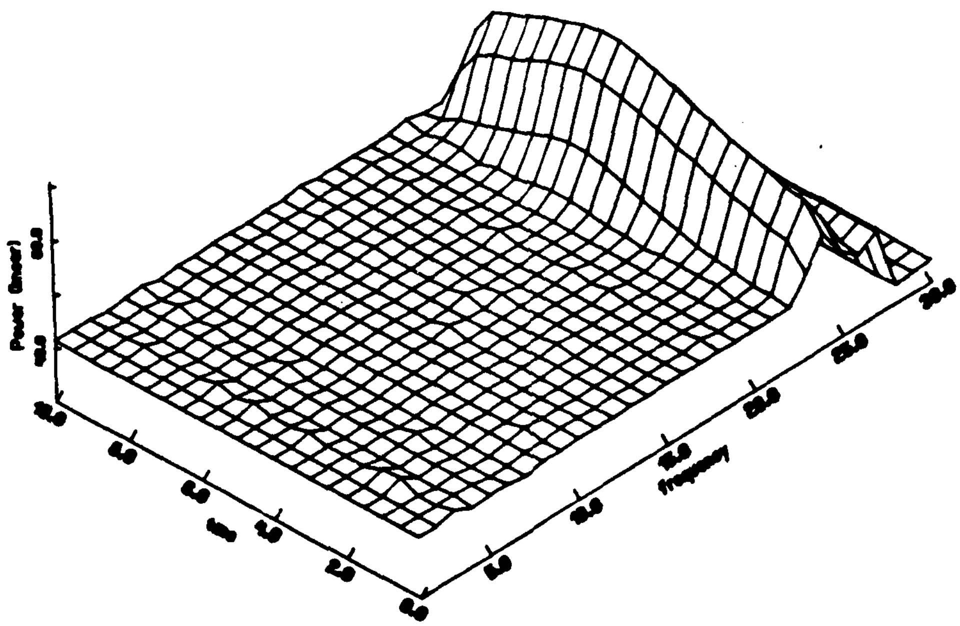


Figure D-10: Isometric View of One-Third Octave Band Spectrum of Signal 10 at Half-Second Intervals.